JYU DISSERTATIONS 71

Liisa Laine

Essays on the Economics of Health Care Markets

JYVÄSKYLÄ UNIVERSITY
SCHOOL OF BUSINESS AND ECONOMICS

JYU DISSERTATIONS 71

Liisa Laine

Essays on the Economics of Health Care Markets

Esitetään Jyväskylän yliopiston kauppakorkeakoulun suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa S212
huhtikuun 26. päivänä 2019 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Jyväskylä University School of Business and Economics,
in building Seminarium, auditorium S212, on April 26, 2019 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2019

Editors
Tuomo Takala
Jyväskylä University School of Business and Economics
Ville Korkiakangas
Open Science Centre, University of Jyväskylä

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7722-1

ISBN 978-951-39-7722-1 (PDF)
URN:ISBN:978-951-39-7722-1
ISSN 2489-9003

ABSTRACT

Laine, Liisa T.
Essays on the Economics of Health Care Markets
Jyväskylä: University of Jyväskylä, 2019, 220 p.
(JYU Dissertations
ISSN 2489-9003;71)
ISBN 978-951-39-7722-1 (PDF)

This doctoral dissertation studies competition in health care markets and the labor market consequences of health behavior. The focus is on examining how the special characteristics of health care markets and the provider incentives affect the market structure, health care qualities, prices, and social welfare. The dissertation consists of an introductory chapter and four separate essays. The introductory chapter discusses the special features of health care markets, research questions, and methods and data used and provides an overview of the main results and policy implications. The first three essays are theoretical contributions and the fourth is empirical.

The first essay studies price and quality competition in markets with public and private providers. We show that equilibrium qualities are often inefficient, but under some conditions on the consumer valuation distribution, equilibrium qualities coincide with the first best.

The second essay extends the analysis to consider qualities with multiple attributes. I show that additional assumptions on the per-unit production cost of quality are required for the equilibrium qualities to be efficient. The results of the first two essays reveal which properties on the consumer preference distribution and the per-unit production costs have been driving the results in the previous literature.

The third essay studies how regulation of health care payment schemes and licensing affect health care providers' entry and health care quality decisions when some patients have inaccurate quality perceptions. I show that entry licensing combined with a regulated prospective payment scheme may be preferable to an unregulated entry and a more complicated provider reimbursement scheme. Providing better information about provider quality may also have different direct and indirect effects depending on whether patients underreact or overreact to health care quality.

The fourth essay analyzes linkages between the different durations of being overweight and long-term labor market outcomes. We find that being persistently overweight in early adulthood drives lower subsequent long-term earnings for women and men. The potential mechanism seems to be different for women and men. For women, the earnings penalty is related to their weaker labor market attachment. For men, the earnings penalty is not related to their labor market attachment and instead is related to something that erodes their earnings power on the labor market throughout their life cycle.

Keywords: Competition, prices, quality, public and private firms, multi-dimensional product differentiation, regulation, entry, mixed payment schemes, overweight, obesity, long-term labor market outcomes, labor market attachment.

TIIVISTELMÄ (ABSTRACT IN FINNISH)

Laine, Liisa T.
Taloustieteellisiä tutkimuksia terveydenhuoltomarkkinoista
Jyväskylä: University of Jyväskylä, 2019, 220 p.
(JYU Dissertations
ISSN 2489-9003; 71)
ISBN 978-951-39-7722-1 (PDF)

Väitöskirjassa tarkastellaan terveydenhuoltomarkkinoiden erityispiirteiden ja terveydenhuoltojärjestelmän ominaisuuksien vaikutuksia palveluntarjoajien hinnoitteluun, niiden tuottamien terveyspalveluiden laatuun, markkinarakenteeseen ja yhteiskunnalliseen hyvinvointiin. Väitöskirja koostuu johdantoluvusta sekä neljästä tutkimuksesta. Kolme ensimmäistä tutkimusta ovat teoreettisia ja neljäs empiirinen.

Ensimmäinen tutkimus tarkastelee laatu- ja hintakilpailun vaikutuksia ja tehokkuutta epätäydellisesti kilpailluilla markkinoilla, joilla tarjottavat tuotteet ovat vertikaalisesti differoituja ja joilla on sekä julkisia että yksityisiä palveluntarjoajia. Tutkimuksen päätulos on, että tiettyjen kuluttajien preferenssijakauman muotoon, julkisyrityksen tavoitefunktioon ja tuotantoteknologiaan liittyvien erityisehtojen vallitessa palveluntarjoajien valitsemat laadut voivat markkinatasapainossa olla yhteiskunnan näkökulmasta optimaalisesti valittuja.

Toinen tutkimus jatkaa samassa aihepiirissä ja olettaa, että tuotteen laatu koostuu useasta ominaisuudesta. Tutkimuksessa löydetään, etteivät ensimmäisessä tutkimuksessa esitetyt erityisehdot ole tällöin riittäviä sosiaalisen optimin saavuttamiseksi. Kahden ensimmäisen tutkimuksen tulokset näyttävät, mitkä kuluttajien preferenssijakauman ja palvelun laaduntuotannon tuotantoteknologian ominaisuudet ovat ajaneet aiemman tutkimuskirjallisuuden tuloksia.

Kolmannessa tutkimuksessa tarkastellaan terveydenhuollon korvausjärjestelmien ja markkinoillepääsyn sääntelyn merkitystä. Tutkimuksessa rakennetaan malli, jossa palveluntarjoajien tuotteet ovat laadullisesti differoituja, palveluntarjoajat kilpailevat asiakkaista laatuvalinnoillaan ja osa potilaista havaitsee laadun epätäydellisesti. Tulosten perusteella markkinoillepääsyn sääntely esimerkiksi toimilupien avulla tietynlaiseen etukäteen määriteltyihin korvauksiin perustuvaan korvausjärjestelmään yhdistettynä voi johtaa yhteiskunnan kannalta optimaaliseen tulemaan.

Neljännessä tutkimuksessa tarkastellaan ylipainoisuuden keston ja pitkän aikavälin ansiotulojen ja työmarkkinoille kiinnittymisen välistä yhteyttä. Tulokset viittaavat, että pysyvä ylipaino on yhteydessä alempiin pitkän aikavälin ansiotuloihin sekä naisilla että miehillä. Mahdolliset mekanismit ovat erilaiset: naisilla alhaisemmat pitkän aikavälin ansiotulot näyttävät liittyvän heikompaan työmarkkinoille kiinnittymiseen läpi elinkaaren, kun taas miehillä tekijät näyttävät liittyvän sellaisiin tutkimusaineistosta havaitsemattomiin tekijöihin, jotka heikentävät yksilöiden ansaintakykyä läpi elinkaaren.

Avainsanat: Kilpailu, hinnat, laatu, julkiset ja yksityiset palveluntarjoajat, moniuloitteinen tuotedifferointi, sääntely, markkinoilletulo, korvausjärjestelmä, ylipaino, lihavuus, pitkän aikavälin ansiotulot, työmarkkinoille kiinnittyminen.

Author Liisa T. Laine
School of Business and Economics
University of Jyväskylä

Supervisor Professor Ari Hyytinen
School of Business and Economics
University of Jyväskylä

Reviewers Professor Ismo Linnosmaa
Department of Health and Social Management
University of Eastern Finland

Senior Adviser Tuomas Takalo
PhD Adjunct Professor
Monetary Policy and Research Department
Bank of Finland

Opponent Senior Adviser Tuomas Takalo
PhD Adjunct Professor
Monetary Policy and Research Department
Bank of Finland

ACKNOWLEDGEMENTS

Working on this dissertation has been a great experience, and I take this opportunity to express my deepest gratitude to several individuals and institutions that made it possible. I want to begin by thanking Yrjö Jahnsson Foundation, Alfred Kordelin Foundation, University of Jyväskylä, OP-Pohjola-ryhmän tutkimussäätiö, KAUTE foundation, and Tutkijat maailmalle for their financial support during my graduate studies and research visits to the U.S. This dissertation would not exist without their generous support.

I would like to thank my supervisor Professor Ari Hyytinen for his support, advice, and his insightful and precise comments. Thanks to Ari, I also discovered the field of health economics. I thank my pre-examiners Professor Ismo Linnosmaa and Senior Adviser Tuomas Takalo for reviewing my dissertation and for their extremely valuable and detailed feedback. It is also an honor to have Senior Adviser Tuomas Takalo as my opponent at the public examination.

I have been very fortunate to work in so many inspirational research environments in the U.S. over the past years. In particular, I am extremely grateful to Professor Ching-to Albert Ma for hosting me during my visit at Boston University. I cannot thank enough for his supervision, guidance, support, and encouraging me to tackle challenging problems. I am also grateful to Francois Gerard for being my host at Columbia University and to Anirban Basu for hosting my stay at University of Washington. I thank Albert Ma, Juan Ortner, and Marc Rysman at Boston University, Josh Angrist, Amy Finkelstein, Jonathan Gruber, and Heidi Williams at MIT, and Katherine Ho at Columbia University for letting me attend their courses.

I owe gratitude to the faculty and staff at JSBE in addition to VATT and HECER for their hospitality and support. I thank Hannu Vartiainen for helping me with arranging my visit to Boston University. I also thank Benjamin Solow, Tuuli Paukkeri, Francesca Valentini, Mikaela Carlström, Suvi Vasama, Heikki Lehkonen, Isa Kuosmanen, Petteri Juvonen, Tanja Saxell, Robin Stitzing, and all other friends at JSBE, Economicum, Boston University, MIT, Columbia University, and University of Washington for the amazing peer support and inspiration.

Special thanks go to my parents Päivi and Mikko. I am forever grateful for their trust and support, for encouraging me to ask questions, to do things my own way and aim high. I also thank my amazing siblings Leena and Hannu and their families, my dear grandparents and all my friends near and far for their support and bringing balance to my life. I adore and admire all of you.

Nothing is better than having the funniest and most intelligent man on your side. Thank you, my husband Miikka for your love. You inspire me every day. I dedicate this dissertation to the apple of my eye, Elma. You have brought such endless joy to our lives. I love you.

Philadelphia, March 2019
Liisa T. Laine

CONTENTS

ABSTRACT
TIIVISTELMÄ (ABSTRACT IN FINNISH)
ACKNOWLEDGEMENTS
CONTENTS

1 INTRODUCTION 11
1.1 Background 11
1.1.1 Research questions 15
1.1.2 Theoretical frameworks, methods, and data 16
1.2 Overview of the essays, results, and contributions 18
1.2.1 Essay 1: Quality and competition between public and private firms 18
1.2.2 Essay 2: Multi-dimensional product differentiation in a mixed oligopoly 19
1.2.3 Essay 3: Health care provider entry and quality competition under regulated payments and inaccurate quality perceptions 21
1.2.4 Essay 4: Temporary and persistent overweight and long-term labor market outcomes 22
1.3 Synthesis of the policy implications 24
References 27

2 QUALITY AND COMPETITION BETWEEN PUBLIC AND PRIVATE FIRMS 32

3 MULTI-DIMENSIONAL PRODUCT DIFFERENTIATION IN A MIXED OLIGOPOLY 53
3.1 Introduction 54
3.2 The model 58
3.2.1 Consumers 58
3.2.2 Public and private firms 59
3.2.3 Extensive form 62
3.3 Allocation, social surplus, and the first best 63
3.3.1 General definition of the first best 63
3.3.2 Quality differentiation in the first best 65
3.3.3 Numerical examples of the first best when quality-valuation density is a step function 65
3.4 Equilibria with low first quality attribute at the public firm 68
3.4.1 Subgame-perfect equilibrium prices 68
3.4.2 Subgame-perfect equilibrium qualities 72
3.4.3 Quality differentiation and efficiency when quality-valuation density is a step function 76
3.4.4 Numerical examples for $q_1 < r_1$ 77

3.5 Equilibria with high first quality attribute at the public firm 80
3.5.1 Subgame-perfect equilibrium prices 80
3.5.2 Subgame-perfect equilibrium qualities............................ 83
3.5.3 Quality differentiation and efficiency when quality-valuation density is a step function 86
3.5.4 Numerical examples for $q_1 > r_1$ 87
3.6 Conclusions 90
3.A Appendix: Proofs and derivations.. 93
3.A.1 Proofs of the results in Section 3.3................................... 93
3.A.2 Proofs of the results in Section 3.4................................... 98
3.A.3 Proofs of the results in Section 3.5................................... 113
3.A.4 Derivations of the equations for the examples 126
3.A.5 Sufficient second order conditions and equilibrium existence.. 134
3.A.6 Details about the numerical simulations 135
References 137

4 HEALTH CARE PROVIDER ENTRY AND QUALITY COMPETITION UNDER REGULATED PAYMENTS AND INACCURATE QUALITY PERCEPTIONS 139
4.1 Introduction.. 140
4.2 The Model 144
4.2.1 Patients.. 144
4.2.2 Providers and the regulator 145
4.2.3 Extensive form 145
4.2.4 Patient behavior 146
4.3 Equilibrium 147
4.3.1 Quality choices.. 147
4.3.2 Entry 150
4.4 Payment scheme regulation.. 152
4.4.1 First best 152
4.4.2 Equilbrium qualities and the number of providers when the number of providers is contractible 154
4.4.3 Equilbrium qualities and the number of providers when the number of providers is not contractible 155
4.5 Policy discussion 157
4.6 Conclusions 158
4.A Appendix: Proofs and derivations.. 160
4.A.1 Proofs of the results in Section 4.3.1 160
4.A.2 Proofs of the results in Section 4.3.2 166
4.A.3 Proofs of the results in Section 4.4.1 168
4.A.4 Proofs of the results in Section 4.4.2 172
4.A.5 Proofs of the results in Section 4.4.3 174
4.A.6 Example 181
References 182

5 TEMPORARY AND PERSISTENT OVERWEIGHT AND LONG-TERM LABOR MARKET OUTCOMES 184
5.1 Introduction 185
5.2 Data and variables 188
5.2.1 Data 188
5.2.2 Variables 189
5.2.3 Summary statistics 191
5.3 Empirical framework 193
5.3.1 Empirical strategy 193
5.3.2 Potential mechanisms 196
5.4 Results 198
5.4.1 Baseline results: OLS regressions 198
5.4.2 The role of family environment 199
5.4.3 The role of genetics 201
5.5 Robustness 204
5.5.1 Alternative BMI-thresholds 204
5.5.2 Alternative characterization of weight history 205
5.5.3 Alternative outcome measures 206
5.5.4 Additional control for physical condition 206
5.6 Discussion 207
5.7 Conclusions 208
5.A Appendix 209
References 217

YHTEENVETO (SUMMARY IN FINNISH) 221

1 INTRODUCTION

This doctoral dissertation consists of three essays on competition in health care markets and one essay on the economics of health behavior. In the first three essays, I apply theoretical industrial organization models to questions related to competition in health care markets. The first two essays focus on price and quality competition and efficiency in markets with public and private providers. My third essay studies how regulation of health care payment schemes and licensing affect health care providers' entry and health care quality decisions when some patients have inaccurate quality perceptions. My fourth essay provides an empirical analysis on the linkages between the different durations of being overweight and long-term labor market outcomes.

The first three essays of this dissertation study questions related to the functioning of health care markets and the interactions between firms. Thus, in addition to health economics, these essays overlap with the field of industrial organization. The fourth essay contributes to health economics and to the field of labor economics. The first and third essays are co-authored with different co-authors, and the second and third essays are single-authored.

1.1 Background

The health care sector constitutes a significant part of the economies of all developed countries throughout the world. For example, the total health care spending in 2016 in Finland amounted to 9.6 percent, Sweden 11 percent, Germany 11.3 percent, France 11 percent, and the U.S. 17.2 percent of the GDP. In Finland that amounted to 20.5 billion (euros), and in the U.S. to 3.3 trillion dollars (THL 2018). Per-capita spending on health care varies across developed countries: for example, the estimated total health care spending per-capita in 2017 was 4,176 dollars in Finland, 5,511 dollars in Sweden, 6,351 in Norway, 5,729 in Germany, and 10,209 in the U.S., while on average per-capita spending of the 35 OECD countries was 4,003 dollars (OECD 2018).[1] In addition to the size of the health care sector in economies, it also has massive consequences for the populations well-being.

The functioning of health care markets plays an important role in the delivery of health care. However, there are several characteristics in health care markets that make them different from other markets.[2] First, the demand for health care depends on the health status of the population, which is difficult to predict because an individual's health status is uncertain. The supply of health care, on the other hand, depends on the availability and efficacy of medical treatments. Because there is uncertainty in both demand and supply of health care, there is a need for health insurance (McGuire 2011).

The prevalence of public and private providers is common in health care markets (Barros and Siciliani 2011). In many countries, health care is at least partly publicly financed. Medical care is often provided jointly by private, public, or other non-profit providers (Besley and Malcomson 2018). Price determination differs from traditional markets. First, prices are often regulated. Second, because health care is usually reimbursed by an insurer, demand is driven by non-price characteristics such as quality. Health care markets are also characterized by the presence of asymmetric information (see the seminal finding by Arrow 1963). For example, patients may not have all information on the effectiveness of treatments or the quality of health care providers, and providers may have limited information about the patients (Chandra et al. 2011; McGuire 2000). All the characteristics described above imply that the market may not be able to assign resources efficiently in the health care sector, which means there is also room for regulation (Dranove and Satterthwaite 2000).

Population aging and the prevalence of chronic diseases put additional challenges on the functioning of health care markets and the efficient use of scarce health care resources in developed countries. Increasing demand for health care also puts pressure on the sustainability of public finance. Because of these challenges, many recent and ongoing health care reforms in several countries are working on finding solutions to improve efficiency of health care markets, provide better health care quality and efficacy, and improve access to health care by increasing provider competition.[3] With a goal of creating a more market-based health care system, many previously purely public systems are opened to private health care provision in addition to allowing patients to choose their preferred medical provider more freely.[4] However, because of the several special features

1 The U.S. constitutes an outlier in terms of health care spending: the spending in the U.S. in 2016 was 3.3 trillion dollars, and in comparison the total GDP in Germany in 2016 was 3.5 trillion dollars (United Nations 2017).

2 Instead of there being a single feature that makes health care markets different from other markets, there are actually several features known from other markets, which combined make health care markets unique (Barros and Martinez-Giralt 2013).

3 For great summaries on the industrial organization of health care markets see Gaynor et al. (2015) and Pauly et al. (2011).

4 The reforms that allow more or free choice of medical providers are often also called patient choice programs. Examples of countries that have implemented health care reforms

of health care markets described above, the supply and demand responses to these policies can be complicated and it is unclear whether the policies actually meet the goals originally set by the policy maker.

These challenges combined with the complexity of health care markets makes the field of health economics relevant. Health economics studies the allocation of resources to and within the health care sector in the economy and the functioning of the health care markets. Because the core questions of health economics focus on the governing rules that provide incentives to health care providers and individuals to use the scarce resources in a way that maximizes social surplus (Barros and Martinez-Giralt 2013), in addition to the academic contributions, the research is also relevant to the policy makers.[5]

This dissertation provides new results on these themes. The first two essays of this dissertation focus on (unregulated) price and quality competition between public and private firms.[6] Simultaneous provision by both public and private providers is common in health care markets. For example, many European health care markets often dominated by public providers, and in other countries such as in the U.S. the private market is more active. Hospital markets are also often consolidated (Cuellar and Gertler 2003). The small number of providers thus calls for modeling firm behavior in a way that takes into account that they are aware of one another's actions. The presence of public and private providers in the market means that the objective function of at least one of the firms differs from that of the others. This means that policy implications derived using the standard economic models which are based on pure private competition may not apply to markets with public and private providers.

The models also should give guidance on the welfare effects of quality segmentation in markets, quality segmentation meaning whether the private or the public provider functions at the lower or higher quality part of the market. Empirically, there is plenty of variation in firms' the quality segmentations in markets: for example, in Europe the quality of care is diverse across public and private hospitals (Kruse et al. n.d.). On the other hand, in the U.S., four out of the five best U.S. hospitals are private (the U.S. news ranking in 2016-2017), whereas the public nursing homes in the U.S. are of a higher quality than private nursing homes (Comondore et al. 2009).

Opening formerly public markets to free entry of providers has been suggested to improve provider efficiency because of increased competition. However, the effects of relaxing entry regulations are unclear. Opening markets for private entry may lead to an inefficient number of providers in the market: there

with some version of patient choice programs are Australia, Belgium, Denmark, Israel, the Netherlands, Sweden, the United Kingdom (UK), and the Medicare Part D program in the United States. For discussions on the patient choice programs in different countries see for example Cooper et al. (2011), Glied and Altman (2017), Gruber (2017), Dietrichson et al. (2016), Santos et al. (2017), and Ikkersheim and Koolman (2012).

5 For example, Barros and Martinez-Giralt (2013) and Dranove and Satterthwaite (2000) provide excellent descriptions on the field of health economics.

6 In industrial organization literature these markets are often called mixed oligopolies (De Fraja and Delbono 1990).

may be either too few (Dixit and Stiglitz 1977; Spence 1976) or too many providers (Perry 1984; Salop 1979; Weizsacker 1980) compared to an efficient number of providers (for papers that consider both under- and over-entry, see for example Mankiw and Whinston (1986) and Amir et al. (2014)). Also, because health care is reimbursed by an insurer, patients' choices are driven by non-price characteristics such as quality, which means that demand response requires that patients have knowledge of what health care quality is. Even if patients have perfect knowledge of quality, providers may react to increased demand differently: for example, there can be over-provision to low-severity patients or under-provision to high-severity patients (skimping) or even avoidance of high-severity patients (dumping) (Ellis 1998).

In reality, information on provider quality may be difficult to obtain and process, and patients may have incorrect information about providers' quality and services. The choice process can be difficult too, as illustrated in the markets for other complex products, such as financial investments and health insurance plans (Abaluck and Adams 2017; Abaluck and Gruber 2011; Handel and Kolstad 2015). My third essay uses a model that combines provider entry and quality choices, non-price competition, quality misperception, and payment regulation, which are all typical characteristics of health care markets and reforms described above.

My fourth essay moves from studying the health care markets and health care provision to understanding the association between the timing of being overweight and long-term labor market outcomes and the mechanisms underlying this association. Being obese and overweight is very common: in the U.S. almost 40 percent of men and 30 percent of women were obese in 2007-2012, and the shares of obese and overweight individuals have been increasing during the past decades (Flegal et al., 2016). Also, in Finland, 75 percent of working aged men and 50 percent of working aged women are overweight, and almost 25 percent of the working aged population is obese (Lundqvist et al. 2018). Being overweight or obese are associated with adverse health outcomes. The previous literature also suggests being overweight or obese are also associated with weaker labor-market outcomes. Despite the importance of the topic, very little is known about the relationship between timing and changes in being overweight and long-term labor market outcomes, and about the mechanisms at work.

Gaining a deeper understanding on the association between dynamics of being overweight and the labor market outcomes is important for policy design. For example, if the earnings penalty depends on how long people have been overweight, policies that aim at reducing discrimination against overweight individuals are likely to have heterogeneous effects, depending on the individuals' weight history. The same applies to policies that seek to help currently overweight individuals by providing them with more information about healthy food options and the health benefits of being physically active. Earlier interventions that aim at prevention of becoming overweight would be beneficial if having been previously overweight is a determinant of the earnings penalty or weaker labor market attachment. For example, the case for early obesity prevention programs and strategies becomes stronger if they can enhance long-term labor market outcomes in addition to the population's health by targeting children and adolescents.

The remainder of this introductory chapter introduces my research questions and describes the theoretical framework and the methods used in this dissertation. Section 1.2 gives an overview of the essays and their contributions. Section 1.3 provides a synthesis of the results of this dissertation and the main policy implications.

1.1.1 Research questions

The first two essays study price and quality competition and efficiency in markets with public and private firms. The third essay studies impacts of health care payment scheme regulation for health care providers' entry and quality competition when some patients have inaccurate quality perceptions. The fourth essay provides an empirical analysis on the linkages between different durations of being overweight and long-term labor market outcomes. I discuss the research questions of each essay next.

The research questions of the first essay are the following. How does the presence of a social-surplus maximizing public firm affect the strategic behavior of a profit maximizing private firm when they both compete on prices and qualities? What are the strategic responses of each firm? How efficient are the qualities and the prices in a mixed duopoly compared to if the goods were supplied by a private duopoly? Does the standard result from private product differentiation, which is that the products are differentiated to relax price competition, hold in the mixed duopoly? How are these results affected by the underlying distribution of consumer preferences?

The second essay continues with the theme of the first essay but extends the model to consider qualities with multiple attributes. Because this is the first essay studying product differentiation in a vertically differentiated mixed oligopoly, a natural question to ask is how to characterize the first best and equilibrium in a quality-price competition mixed duopoly model when a product quality consists of multiple quality attributes? How this new assumption affects the results in the first essay, and if so, how? Will there be differentiation in all quality attributes in the equilibrium?

The third essay investigates the effects of the provider payment scheme regulation for health care provider entry decisions and health care quality choices. Can the equilibrium outcome be effcient? Do the effects of payment scheme regulation differ if the provider entry is contractible, for example by giving licenses, compared to a situation of free provider entry? How is this affected by the competition environment in the market and patients' misperceptions of health care quality?

Finally, the research questions of the fourth essay are: how are different durations of being overweight associated with long-term earnings and employment? What is the relationship between the timing of and changes in being overweight and long-term labor market outcomes? What are the mechanisms underlying this association?

1.1.2 Theoretical frameworks, methods, and data

The essays of this dissertation cover a wide range of topics and therefore there exists no common specific theoretical nor empirical literature to which the they contribute. However, in each essay below I discuss the linkages with and contributions to the existing research literature in depth.

The first three essays are theoretical and use multistage games with observed actions. These models are used for studying the strategic behavior and interaction of the firms in different market settings and thus are also very suitable for analyzing strategic interaction between providers in health care markets.

I use models of imperfect competition and product differentiation. The first essay uses a standard duopoly model of vertical product differentiation, originating from Gabszewicz and Thisse (1979) and Shaked and Sutton (1982, 1983). The only way we deviate from the standard textbook model of vertical product differentiation is that we assume one firm to be a social-surplus maximizing public firm, whereas the other firm is a profit-maximizing private firm. The second essay also makes less restrictive assumptions on the distribution that represents consumer valuations and on the variable, unit production cost of the good than the model used in the previous literature. My second essay builds on the model used in the first essay but extends the model by assuming that instead of a single quality attribute, product quality consists of two quality attributes.[7] Because public and private providers are common in many other markets, such as education, transportation, and utility, these models can be applied more broadly to many other markets as well.

The first two essays study strategic interaction between firms with different objective functions in unregulated markets. My third essay considers two new distinct features of health care markets: regulation and non-price competition. In this essay, I use the model on localized competition of multiple firms by Salop (1979), which allows me to focus on provider entry. I make the following extensions to the standard model. First, instead of a price that is usually the choice variable in the standard Salop model, I let the providers choose the quality of the medical good or service they offer. Second, I assume that the regulator fully covers patients' medical expenses, and the only cost patients face is a transportation cost to the provider's location. Third, I assume that the regulator makes a choice on how the payment scheme is regulated. Fourth, I let some patients overreact or underreact to perceived health care quality, which I call quality misperception.[8] As far as I am aware of, this is the first paper that combines all these features of health care markets in one model.

The first three essays are theoretical contributions to the current economic

7 For the previous papers studying multi-dimensional product differentiation in private duopolies see, for example, Barigozzi and Ma 2018, Lauga and Ofek (2011),Vandenbosch and Weinberg (1995), and Irmen and Thisse (1998).

8 This terminology comes from Kahneman et al. (1997)

literature. Together with empirical research, findings from theoretical models can be used in guiding economic decision making. One of the key benefits is that economic theory can be used in obtaining insights on the functioning of the market. In particular, despite of being simple representations of health care markets, models can be used in to pin down what market parameters are important and how these parameters could be measured, giving a basis for quantification and welfare calculation.

My fourth essay is an empirical analysis on the association between individual health behavior (overweight) and long-term labor market outcomes. The data used in this essay combines two unique and detailed data sets. First is the Finnish Longitudinal Employer-Employee Data (FLEED), which is an annual panel that covers the working-age population in Finland over the period from 1990 to 2009. These data are used to create our measures for long-term labor market outcomes. The second data set is the Older Finnish Twin Cohort Study, which was launched in 1974, with the survey being carried out in 1975, 1981, and 1990. The Older Finnish Twin Cohort Study is a survey that covers all Finnish twin pairs born before 1958 and for which both co-twins were alive in 1975. This data set is used to derive measures for Body Mass Index (BMI), which is used to construct our main regressors of interest.

The empirical strategy in my fourth essay exploits within-twin variation to control for several key sources of omitted variable bias, mainly the family environment and genes shared by twins. Controlling for family environment is important because the shared home environment affects weight development (Gregory and Ruhm 2011; Segal and Allison 2002), and there are family-level peer effects (Gwozdz et al. 2015). Controlling for genetic variation by focusing solely on identical twins is motivated by the previous literature, which has documented that the genetically heritable component of weight is very high (Cawley 2004, 2015; Farooqi and O'Rahilly 2007). By focusing on identical twins thus removes the genetic effects. This also controls for differences in appearance.

An obvious challenge in the literature studying the economic consequences of being overweight and obese is to find plausibly exogenous variation in weight that allows one to identify and estimate the causal effect of being overweight or obese. Previous papers have used various instruments when trying to solve this problem: sibling-instruments (Cawley 2004; Cawley and Meyerhoefer 2012; Kline and Tobias 2008; Lindeboom et al. 2010) lagged-weights (Averett and Korenman 1996; Baum and Ford 2004; Cawley 2004; Sargent and Blanchflower 1994), and the prevalence of obesity (Morris 2007).

One of the more current approaches is to use genetic risk scores as an instrument (Böckerman et al. 2016). While we find the approach of using genetic risk scores valuable, this approach cannot be used to address the specific research question that we are interested in. As far as we are aware, medical science and behavioral genetics have not identified genes that would predict individuals' weight in different parts of their life-cycle or, relatedly, the duration of being overweight. Second, even if such instruments became available, it is not clear that they would satisfy the required exclusion restriction for instrumental variable estimation in our context.

1.2 Overview of the essays, results, and contributions

1.2.1 Essay 1: Quality and competition between public and private firms

In many markets such as health care, education, transportation, and utility, public and private firms jointly serve consumers.[9] Product quality is also a major concern in these markets. The concern stems from a fundamental point made by Spence (1975). The social benefit of quality is the sum of consumers' valuations because a good's quality benefits all buyers. At the social optimum, the average consumer valuation of quality should be equal to the marginal cost of quality. Yet, because a profit-maximizing firm is only concerned with the consumer who is indifferent between buying and not buying, a firm's choice of quality will be one that equates this marginal consumer's valuation to the marginal cost of quality. This gives the classic Spence (1975) result: even when products are priced at marginal costs, their qualities will be inefficient.

My first essay studies whether a market in which a social-surplus maximizing public and profit-maximizing private firms compete may be a mechanism for remedying this inefficiency. We use a standard model of vertical product differentiation originating from Gabszewicz and Thisse (1979) and Shaked and Sutton (1982, 1983). In the first stage of the model, two firms simultaneously choose product qualities. In the second stage, firms simultaneously pick their product prices, and then consumers pick their preferred provider. Consumers' quality valuations follow a general distribution for the quality valuation, and the two firms have access to the same technology. The only difference from the standard model of product differentiation is that one firm (public) maximizes social surplus and one firm (private) maximizes profits.[10]

We find that the game has multiple equilibria: in one equilibrium class the public firm's product quality is lower than the private firm's quality, and in others, the opposite is true. Second, after characterizing the subgame-perfect equilibria we show that equilibrium qualities are often inefficient, but under some conditions on the consumer valuation distribution, equilibrium qualities coincide with the first best. These results have several implications for competition

[9] For recent surveys on public and private provision in health care and education, see Barros and Siciliani (2011) and Urquiola (2016).

[10] This objective function makes our model different from the literature studying not-for-profit providers, such as recent papers by Besley and Malcomson (2018), who study market entry by for-profit providers, when quality is unobserved and the incumbent is non-profit, and Stenbacka and Tombak (2018), who study reimbursement policies and coverage in the health care markets with a profit-maximizing private firm and a non-profit firm. In Besley and Malcomson (2018) the not-for-profit have a stronger preference towards quality relative to for-profits, whereas in Stenbacka and Tombak (2018) the for-profit firm maximizes consumer surplus.

policy.

We present the general (sufficient) conditions on consumers' quality-valuation distribution under which qualities in low-public-quality equilibria are efficient, as well as general conditions under which qualities in high-public-quality equilibria are efficient. We show that in equilibria where the public firm produces at a low quality, equilibrium qualities are first best when the valuation distribution has a linear inverse hazard rate. In equilibria where the public firm produces at a high quality, equilibrium qualities are first best when the valuation distribution has a linear inverse reverse hazard rate. In general, because linear inverse hazard and inverse reverse hazard rates are special cases in the broad class of valuation distributions, a generic valuation distribution violates linearity. In an inefficient equilibrium, both firms' qualities are either too high or too low relative to the first best.

We contribute to the previous literature in several ways. By constructing the two equilibrium classes we illustrate the variety of quality segmentations in the markets mentioned above. More importantly, we show that in these markets equilibrium qualities can sometimes be efficient but are often inefficient. This finding is in a stark contrast to the findings from the previous mixed oligopoly literature that has suggested that having a social-surplus maximizing public firm causes the equilibrium qualities to be the socially optimal (Cremer et al. 1991; Grilo 1994). Our results also simultaneously reveal the limitation of the uniform distribution as well as which properties of the uniform distribution (linear inverse hazard and linear inverse reverse hazard rates) have been the driver of earlier results. Furthermore, when consumer valuations follow a uniform distribution, the issue of multiple equilibria is moot for a duopoly. By contrast, we show that multiple equilibria are important for general distributions.

1.2.2 Essay 2: Multi-dimensional product differentiation in a mixed oligopoly

My second essay continues with the theme of public-private competition, but it extends the model of the first essay to allow product quality to consist of two attributes. In markets where simultaneous provision by public and private providers is common, product quality also often consists of multiple attributes. For example, the choice of a non-urgent care or a long-term care provider is made based on different types of quality attributes such as the (perceived) clinical quality, the amenities, the waiting times, and the geographic location of the health care provider.

When product quality has multiple attributes, an additional question to ask in addition to efficiency is whether firms in a mixed oligopoly differentiate in all dimensions in equilibrium. In private oligopolies, firms use (excess) product differentiation to relax price competition (Gabszewicz and Thisse 1979; Shaked and Sutton 1982, 1983). If a product quality consists of multiple attributes, private firms do not necessarily differentiate in all of the quality attributes (Barigozzi and Ma 2018). In mixed oligopolies, the competition effect is different because the price of the public firm works as an instrument with the goal to induce or improve market efficiency (Cremer et al. 1991; Grilo 1994, and the results from my first essay).

The model I develop for my second essay follows closely the multi-attribute quality private duopoly model of vertical differentiation by Barigozzi and Ma (2018) and the single-attribute quality mixed duopoly model of my first essay. Similarly to the first essay, I characterize two classes of equilibria. Now, in the first equilibrium class, the private firm provides higher quality in the first quality attribute, and the public firm provides lower quality in the first quality attribute. The second class of equilibrium considers the opposite. For example, the first equilibrium class represents the equilibrium where a public firm provides a more modest quality on amenities than the private firm. In the second equilibrium class a public firm provides higher quality amenities than the private firm. As I will illustrate below, distinguishing between these two equilibrium classes is important because, as illustrated in my first essay, the equilibria may not be isomorphic, i.e. yielding different equilibrium qualities and different social surplus.

I show that if the per-unit cost of producing two quality attributes is separable in the two quality attributes and the consumer valuations are uniform, the equilibrium outcome coincides with the first best. This means that the result of my first essay on the equilibrium qualities being efficient may apply to models with more than one quality attribute. However, I also show that if the per-unit cost of producing two quality attributes is non-separable in the two quality attributes, firms differentiate in both quality attributes. In particular, I give numerical examples using uniform quality-valuation distributions and a non-separable per-unit cost of quality and show that the equilibrium is not efficient. This proof by counterexample shows that unlike in the single-attribute quality mixed duopoly model, linearity in the inverse hazard or inverse reverse hazard are not sufficient conditions for efficiency. My finding that firms differentiate in a two-attribute quality mixed duopoly is in contrast to Rosenman and Munoz-Garcia (2017) who suggest that firms choose not to differentiate in both quality dimensions.

As far as I am aware of, Rosenman and Munoz-Garcia (2017) is the only paper studying two-attribute quality competition between firms with non-symmetric objectives. However, although they study competition between a market share maximizing firm and a profit maximizing firm with horizontal differentiation, they do not consider competition between a social surplus maximizing firm and a private firm and assume that the marginal unit production cost of quality is zero and that the consumer valuations are uniformly distributed. As illustrated by my results, these simplifying assumptions play an important role for the differentiation and efficiency results. Importantly, my results are also in line with the results of Barigozzi and Ma (2018), who show that in a private duopoly with multi-attribute qualities, firms differentiate in multiple dimensions when the per-unit cost of quality is non-separable. This is in contrast to the findings suggested by the previous literature studying multi-dimensional product differentiation.

I also compare the inefficient equilibrium outcomes in the two classes of equilibria to the first best outcomes. These results show that when the equilibrium outcomes are inefficient, the equilibrium qualities and the social surplus can be different in the two equilibrium classes depending on whether the public firm offers the high first quality attribute or the low first quality attribute. Similarly to the results of my first essay, these examples illustrate the importance of considering the different equilibrium classes separately.

1.2.3 Essay 3: Health care provider entry and quality competition under regulated payments and inaccurate quality perceptions

The common goal of recent and ongoing health care reforms is to promote efficiency of health care markets and to improve health care quality by increasing provider competition. As part of these reforms, entry barriers are removed to enhance provider entry, and patients are allowed to choose their preferred medical provider more freely. However, the supply and demand responses to these policies can be complicated. For example free entry may yet lead to the entry of an inefficient number of providers (see for example Mankiw and Whinston 1986; Amir et al. 2014).

Regarding the effects of competition on health care quality, even in competitive private markets with goods priced at marginal costs, quality is usually inefficient (Spence 1975). When health care is reimbursed by an insurer, patients' choice is driven by non-price characteristics such as quality. Thus, a demand response requires that patients know what health care quality is. This can be complicated because information on provider quality may be difficult to obtain and process. Because health care quality is a complex concept, patients may have incorrect information about providers' quality and services. The choice procedure can be difficult as well, as illustrated in the markets for other complex products, such as financial investments and health insurance plans (Abaluck and Adams 2017; Abaluck and Gruber 2011; Handel and Kolstad 2015).

My third essay studies the effects of the provider payment scheme regulation on health care provider entry and health care quality. I study a model in which locally competitive health care providers make strategic choices on entry and health care quality. I build on the model by Salop (1979) with the following extensions. Instead of a price, the providers choose the quality of the medical good or service that they offer. Patients' medical expenses are fully covered by a regulator (an insurer) who decides how these payments are regulated, and the only cost patients face is a transportation cost to the provider's location. Last, some patients may overreact or underreact to perceived health care quality, which I call quality misperception (the terminology follows Kahneman et al. 1997).

I use this model to study two regulatory frameworks. In the first framework the number of providers is contractible, that is the regulator can control the number of providers in the market, and the regulator uses prospective payment as its payment instrument. In the second regulatory framework the number of providers is not contractible, and the regulator uses a mixed payment instrument, a combination of prospective payment and cost reimbursement, as its payment scheme. If the regulator uses prospective payment, it pays a fixed price for delivering a medical care service irrespective of resources used. Diagnosis-related group payments (DRG-payments) are one example of prospective payments. DRG-payments are used in many countries in Europe and in addition to Medicare in the U.S. Capitation payments are another type of prospective payment. If a cost reimbursement is used, the health care provider receives some revenue corresponding to resources that were used to produce medical care (for example fee-for-service).

I find that that entry regulation combined with pure prospective payment allows the regulator to implement the first best quality. This is a result of the following: when the number of providers is regulated, the regulator fixes the number of entering providers to the first best level and chooses the amount of the prospective payment such that it implements the first best quality.

To my knowledge, my paper is the first paper to document this. The result is similar to the classic finding on provider payment scheme regulation by Ma (1994): the contractible degree of horizontal differentiation combined with a pure prospective payment scheme leads to the first best quality. Because Ma (1994) focuses on imperfectly competitive markets and does not consider entry, my result complements the existing literature with the finding that in the locally competitive markets the first best outcome can be obtained by a pure provider payment combined with entry licensing.

When provider entry is not contractible, a mixed reimbursement scheme is needed to implement the first best. In this case, the regulator tries to choose the level of the prospective payment and the cost reimbursement such that the equilibrium number of providers and quality would be efficient. However, the regime in which the regulator is constrained to choosing its payment scheme depends on the exogenous parameters of the model: the competition environment (transportation costs and entry costs) and the (variable) costs of producing quality. Thus, if the regulator is constrained to choosing its payment scheme in a regime where the cost reimbursement is either too high or too low, the equilibrium outcome can be inefficient even when a mixed reimbursement is used.

These results are similar to Bardey et al. (2012), who find that a mixed reimbursement scheme is welfare improving in imperfectly competitive markets with horizontally and vertically differentiated products. My paper complements the findings in Bardey et al. (2012) by showing that in the equilibrium a mixed payment scheme could be used to yield first best entry and quality outcomes also in locally competitive markets.

1.2.4 Essay 4: Temporary and persistent overweight and long-term labor market outcomes

The upward trends in obesity and being overweight have two major economic consequences.First, obesity raises medical costs (Cawley and Meyerhoefer 2012). Second, being overweight may affect an individual's labor market outcomes, such as earnings and employment (Averett 2011; Cawley 2015). My fourth essay describes how different durations of being overweight are associated with long-term earnings and employment. Despite the importance of the topic, very little is known about the relationship between timing and changes in being overweight and long-term labor market outcomes. Even less is known about the mechanisms at work.

We use detailed register-based data on a large sample of fraternal and identical twins born and raised in the same household. Our empirical strategy exploits the within-twin variation to control for several key sources of omitted variable bias, such as the family environment, genes shared by twins (because a shared home environment affects weight development (Gregory and Ruhm 2011; Segal and Allison 2002), and family-level peer effects (Gwozdz et al. 2015)). In addition, we estimate models that fully control for genetic variation in addition to differences in appearance by focusing solely on identical twins because previous literature has documented that the genetically heritable component of weight is very high (Cawley 2004, 2015; Farooqi and O'Rahilly 2007).

We find that being persistently overweight in early adulthood is associated with lower subsequent long-term earnings for women and men. For men, this association is observed only when data on identical twins are used and genetic differences are fully controlled for. Second, we also find that the mechanisms underlying the earnings penalties of being persistently overweight are different for women and men. Persistently overweight women have weaker labor market attachment, working annually almost a month less. The corresponding relationship for persistently overweight men is not statistically significant. Instead, persistently overweight men have lower monthly earnings.

These findings indicate that genetic factors affect the association of being overweight and labor market outcomes for men but not for women. For women, the results are relatively robust across all specifications. For men, they are not: persistently overweight men have smaller long-term earnings and lower monthly earnings after family environment and genetics are simultaneously controlled for.

This suggests an explanation for the lack of consensus in the previous literature regarding the earnings penalty of overweight men (Averett and Korenman 1996; Baum and Ford 2004; Cawley 2004; Sargent and Blanchflower 1994) and, hence, for why the evidence for men has been more mixed (Cawley 2015). It seems that whether and how genetic factors are a source of omitted variable bias differ across women and men. This interpretation is consistent with the view that genetically inherited traits, such as appearance, and other personality traits that are affected by genetics and environment, such as impatience, contribute to weight gains and shape an individual's labor market outcomes.

The previous papers that have studied whether the timing of being overweight matters for labor market outcomes are Han et al. (2011), Chen (2012), and Pinkston (2017). Our study differs from these papers in several ways. Our econometric analysis exploits unique twin data that include longitudinal information both on labor market outcomes and weight measures. The previous literature has only used cross-sectional measures for outcome(s), weight measures, or both. We focus instead on much longer-term labor market outcomes and because it allows us to capture the cumulative labor market effects of being overweight.

Our measurement for the timing of being overweight also differs from Han et al. (2011), Chen (2012), and Pinkston (2017). In Han et al. (2011), weight is measured using data on BMI in the late teenage years and current BMI, whereas in Chen (2012) and Pinkston (2017) weight is measured in early or late adulthood. We instead use three different weight measurements over a period that covers 15 years. The combination of the three long-term outcome measures and access to twin data makes our analysis unique. This combination also provides us with a way of studying the mechanisms behind the weight penalty. In addition to being able to study the link between previous and current weight on long-term labor market outcomes, our approach allows us to study the role of being persistently overweight. To the best of our knowledge this is the first paper to study the link between persistent overweight and labor market outcomes.

Unlike previous studies (Han et al. 2011; Chen 2012; Pinkston 2017), we are able to show that what matters for long-term labor market outcomes is being chronically overweight, that the mechanisms at work differ by gender, and that the association of being chronically overweight with adverse long-term outcomes is not driven by genetic differences. If being overweight reduces people's lifetime earnings as compared to it only reducing current earnings, the implications for lifetime consumption possibilities and for health policy are quite different.

Also, the findings indicate that the mechanism generating the negative relation between being persistently overweight and lower subsequent long-term earnings may be gender-specific. For women, the negative relationship between being persistently overweight and long-term earnings seems to be related to their weaker labor market attachment, which suggests that the mechanism is likely to be related to something that erodes women's labor market attachment throughout the life cycle, for example labor supply decisions or fertility. This result adds an additional perspective to policies that affect the availability of child care and parental leave policies.

Our results also suggest that for persistently overweight men the mechanism seems to be related to something that erodes their earnings power on the labor market but not their labor market attachment throughout their life cycle. This could reflect an unobservable characteristic, for example high discount rate, or characteristics related to skill formation earlier in life, such as underdeveloped social skills or lower self-esteem. Policies that target a reduction of social marginalization may be helpful. On overall, understanding the mechanisms more further seems like a fruitful theme for further research.

1.3 Synthesis of the policy implications

The first three essays of this dissertation study the effects of competition in health care markets. Barros et al. (2016) define competition as "*an instrument to encourage organizations to be more efficient and responsive to the preferences of those they serve.*" For example, privatization is often a policy topic in markets with public provision (Ishibashi and Kaneko 2008; Matsumura 1998), and the goal of privatization has been to spur efficiency in the market. Because the outcomes of competition always depend on the environment and other institutional details, in addition to other policy instruments such as regulation, it is uncertain what kind of incentives competition creates and whether it succeeds in achieving its goals. In sum, whether competition as a policy instrument is more effective than the alternatives (such as regulation) in achieving societal goals is ambiguous.[11]

The findings of this dissertation suggest several policy implications. My first two essays conclude that from the quality perspective joint provision by public firm and private firms may be preferred to pure private provision. This relates to the discussion of privatization of public services: from the quality efficiency perspective, pure privatization is less ideal. The second policy implication considers the objective of the public provider: the findings from the first two essays regard the public firms social-surplus preferences as a normative recommendation; otherwise, the market outcome will be inefficient. The third policy implication highlights the importance of the public provider operating in the correct market segment. According to my results, the welfare consequences of whether the public provider operates in the higher or lower (first attribute) quality segment depends on the distribution of the consumer preferences and the quality production cost technology.

My third essay suggests that entry licensing combined with prospective payment scheme may be a preferred instrument to unregulated entry and a more complicated provider reimbursement scheme. This means that reforms planning on opening markets for the free entry of new providers should consider provider licensing instead of free entry combined with mixed payment scheme regulation. The findings also highlight the role of health care accurate quality information available to patients because quality misperception changes the effectiveness of the regulatory instruments and competition. Also, policy makers should be aware that implications of providing more information about the provider quality may have different direct and indirect effects depending on whether patients underreact or overreact to health care quality.

Overall, my results from the theoretical models show a rich set of potential implications of increasing provider competition in health care markets. If anything, these results suggest that there are many open research questions at the intersection of health economics and industrial organization. In particular, we do not yet understand comprehensively enough how competition in health care markets with all the varying institutional features work. Since competition often increases the role of private providers in the markets, more theoretical and empirical research is needed to understand the effects of competition between public (or non-profit) and private providers, the roles are of public-private partnerships

---

11 In addition to Gaynor et al. (2015) and Pauly et al. (2011) and other summaries on health care competition mentioned in Section 1.1, also Barros et al. (2016) and the report by the Expert Panel on Effective Ways of Investing in Health the European Commission 2015 provide excellent overviews on the competition among health care providers, especially from the European context.

in health care, and different ways of implementing public procurement processes. Because many of the reforms often consider increasing choice for patients too, there is more work to be done in understanding different ways of implementing choice in health care, and what kind of intended and unintended consequences this may have.

References

Abaluck, J. and Adams, A. 2017. What do consumers consider before they choose? Identification from asymmetric demand responses. NBER Working Papers 23566.

Abaluck, J. and Gruber, J. 2011. Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program. The American Economic Review 101, 1180–1210.

Amir, R., De Castro, L., and Koutsougeras, L. 2014. Free entry versus socially optimal entry. Journal of Economic Theory 154, 112–125.

Arrow, K. J. 1963. Uncertainty and the Welfare Economics of Medical Care. The American Economic Review 53 (5), 941–973.

Averett, S. 2011. Labor market consequences: Employment, wages, disability and absenteeism. In The Oxford Handbook of the Social Science of Obesity. Ed. by Cawley, J. Vol. 1. Oxford Handbooks. Oxford University Press, 531–552.

Averett, S. and Korenman, S. 1996. The economic reality of beauty myth. Journal of Human Resources 31 (2), 304–330.

Bardey, D., Canta, C., and Lozachmeur, J.-M. 2012. The regulation of health care providers' payments when horizontal and vertical differentiation matter. Journal of Health Economics 31 (5), 691–704.

Barigozzi, F. and Ma, C.-t. A. 2018. Product differentiation with multiple qualities. International Journal of Industrial Organization (Forthcoming).

Barros, P., Brouwer, W., Thomson, S., and Varkevisser, M. 2016. Competition among health care providers: Helpful or harmful? European Journal of Health Economics 17, 229–233.

Barros, P. and Martinez-Giralt, X. 2013. Health economics: An industrial organization perspective. Routledge.

Barros, P. and Siciliani, L. 2011. Public and private sector interface. In vol. 2. Handbook of Health Economics. Elsevier. Chap. 0, 927–1001.

Baum, C. and Ford, W. 2004. The wage effects of obesity: A longitudinal study. Health Economics 13, 885–899.

Besley, T. and Malcomson, J. M. 2018. Competition in public service provision: The role of not-for-profit providers. Journal of Public Economics 162, 158–172.

Böckerman, P., Cawley, J., Viinikainen, J., Lehtimäki, T., Rovio, S., Seppälä, I., Pehkonen, J., and Raitakari, O. 2018. The effect of weight on labor market outcomes: An application of genetic instrumental variables. Health Economics.

Cawley, J. 2004. The impact of obesity on wages. The Journal of Human Resources 39 (2), 451–474.

Cawley, J. 2015. An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. Journal of Health Economics 43, 244–268.

Cawley, J. and Meyerhoefer, C. 2012. The medical care costs of obesity: An instrumental variables approach. Journal of Health Economics 31 (1), 219–230.

Chandra, A., Cutler, D., and Song, Z. 2011. Who ordered that? The economics of treatment choices in medical care. In Handbook of Health Economics. Ed. by Pauly, M. V., McGuire, T. G., and Barros, P. P. Vol. 2. Handbook of Health Economics. Elsevier, 397–432.

Chen, A. J. 2012. When does weight matter most? Journal of Health Economics 31 (1), 285–295.

Comondore, V. R. et al. 2009. Quality of care in for-profit and not-for-profit nursing homes: systematic review and meta-analysis. BMJ 339.

Cooper, Z., Gibbons, S., Jones, S., and McGuire, A. 2011. Does hospital competition save lives? Evidence from the English NHS patient choice reforms. The Economic Journal 121 (3), 228–260.

Cremer, H., Marchand, M., and Thisse, J.-F. 1991. Mixed oligopoly with differentiated products. International Journal of Industrial Organization 9 (1), 43–53.

Cuellar, A. E. and Gertler, P. J. 2003. Trends in hospital consolidation: The formation of local systems. Health Affairs 22 (6), 77–87.

De Fraja, G. and Delbono, F. 1990. Game theoretic models of mixed oligopoly. Journal of Economic Surveys 4 (1), 1–17.

Dietrichson, J., Ellegard, L. M., and Kjellsson, G. 2016. Patient choice, entry, and the quality of primary care: Evidence from Swedish reforms. Working Papers 36. Lund University, Department of Economics.

Dixit, A. K. and Stiglitz, J. E. 1977. Monopolistic competition and optimum product diversity. The American Economic Review 67 (3), 297–308.

Dranove, D. and Satterthwaite, M. A. 2000. The industrial organization of health care markets. In vol. 1. Handbook of Health Economics. Elsevier, 1093–1139.

Ellis, R. 1998. Creaming, skimping and dumping: provider competition on the intensive and extensive margins. Journal of Health Economics 17 (5), 537–555.

EXPH 2015. Investigating policy options regarding competition among providers of health care services in EU Member States. doi:10.2875/867590.

Farooqi, I. S. and O'Rahilly, S. 2007. Genetic factors in human obesity. Obesity Reviews 8, 37–40.

Flegal, K., Kruszon-Moran, D., Carroll, M., Fryar, C., and Ogden, C. 2016. Trends in obesity among adults in the United States, 2005 to 2014. JAMA 315 (21), 2284–2291.

Gabszewicz, J. J. and Thisse, J.-F. 1979. Price competition, quality and income disparities. Journal of Economic Theory 20 (3), 340–359.

Gaynor, M., Ho, K., and Town, R. J. 2015. The industrial organization of health-care markets. Journal of Economic Literature 53 (2), 235–284.

Glied, S. A. and Altman, S. H. 2017. Beyond antitrust: Health care and health insurance market trends and the future of competition. Health Affairs 36 (9), 1572–1577.

Gregory, C. A. and Ruhm, C. J. 2011. Where does the wage penalty bite? In Economic Aspects of Obesity. University of Chicago Press, 315–347.

Grilo, I. 1994. Mixed Duopoly under Vertical Differentiation. Annales d'Economie et de Statistique (33), 91–112.

Gruber, J. 2017. Delivering public health insurance through private plan choice in the United States. Journal of Economic Perspectives 31 (4), 3–22.

Gwozdz, W., Sousa-Poza, A., Reisch, L. A., Bammann, K., Eiben, G., Kourides, Y., Kovacs, E., Lauria, F., Konstabel, K., Santaliestra-Pasias, A. M., Vyncke, K., and Pigeot, I. 2015. Peer effects on obesity in a sample of European children. IZA Discussion Papers 9051. Institute for the Study of Labor (IZA).

Handel, B. R. and Kolstad, J. T. 2015. Health insurance for "humans": Information frictions, plan choice, and consumer welfare. American Economic Review 105 (8), 2449–2500.

Ikkersheim, D. E. and Koolman, X. 2012. Dutch healthcare reform: Did it result in better patient experiences in hospitals? A comparison of the consumer quality index over time. BMC Health Services Research 12 (1), 76.

Irmen, A. and Thisse, J.-F. 1998. Competition in multi-characteristics spaces: Hotelling was almost right. Journal of Economic Theory 78 (1), 76–102.

Ishibashi, K. and Kaneko, T. 2008. Partial privatization in mixed duopoly with price and quality competition. Journal of Economics 95 (3), 213–231.

Kahneman, D., Wakker, P. P., and Sarin, R. 1997. Back to Bentham? Explorations of experienced utility. The Quarterly Journal of Economics 112 (2), 375–405.

Kline, B. and Tobias, J. L. 2008. The wages of BMI: Bayesian analysis of a skewed treatment-response model with nonparametric endogeneity. Journal of Applied Econometrics 23 (6), 767–793.

Kruse, F. M., Stadhouders, N. W., Adang, E. M., Groenewoud, S., and Jeurissen, P. P. n.d. Do private hospitals outperform public hospitals regarding efficiency, accessibility, and quality of care in the European Union? A literature review. The International Journal of Health Planning and Management 33 (2), 434–453.

Lauga, D. and Ofek, E. 2011. Product positioning in a two-dimensional vertical differentiation model: The role of quality costs. Marketing Schience 30 (5), 903–923.

Lindeboom, M., Lundborg, P., and Klaauw, B. van der 2010. Assessing the impact of obesity on labor market outcomes. Economics & Human Biology 8 (3), 309–319.

Lundqvist, A., Männistö, S., Jousilahti, P., Kaartinen, N., Mäki, P., and Borodulin, K. 2018. Lihavuus. In vol. 1, 45–48.

Ma, C.-t. A. 1994. Health care payment systems: Cost and quality incentives. Journal of Economics & Management Strategy 3 (1), 93–112.

Mankiw, N. G. and Whinston, M. D. 1986. Free entry and social inefficiency. The RAND Journal of Economics 17 (1), 48–58.

Matsumura, T. 1998. Partial privatization in mixed duopoly. Journal of Public Economics 70 (3), 473–483.

McGuire, T. 2000. Physician agency. Handbook of Health Economics 1. Handbook of Health Economics, 461–536.

McGuire, T. 2011. Chapter Five - Demand for Health Insurance. In Handbook of Health Economics. Ed. by Pauly, M. V., McGuire, T. G., and Barros, P. P. Vol. 2. Handbook of Health Economics. Elsevier, 317–396.

Morris, S. 2007. The impact of obesity on employment. Labour Economics 14 (3), 413–433.

OECD 2018. Health spending (indicator). [Visited on 2018-08-13].

Competition in health care markets 2011. In Handbook of Health Economics. Ed. by Pauly, M. V., McGuire, T. G., and Barros, P. P. Vol. 2. Handbook of Health Economics. Elsevier, 499–637.

Perry, M. K. 1984. Scale economies, imperfect competition, and public policy. The Journal of Industrial Economics 32 (3), 313–333.

Pinkston, J. C. 2017. The dynamic effects of obesity on the wages of young workers. Economics & Human Biology 27, 154–166.

Rosenman, R. and Munoz-Garcia, F. 2017. Nonprofit product differentiation. Economics Bulletin 37 (3), 1415–1420.

Salop, S. C. 1979. Monopolistic Competition with Outside Goods. The Bell Journal of Economics 10 (1), 141–156.

Santos, R., Gravelle, H., and Propper, C. 2017. Does Quality Affect Patients' Choice of Doctor? Evidence from England. The Economic Journal 127 (600), 445–494.

Sargent, J. and Blanchflower, D. 1994. Obesity and stature in adolescence and earnings in young adulthood: Analysis of a British birth cohort. Archives of Pediatrics & Adolescent Medicine 148 (7), 681–687.

Segal, N. L. and Allison, D. B. 2002. Twins and virtual twins: Bases of relative body weight revisited. International Journal of Obesity and Related Disorders 26 (4), 437–441.

Shaked, A. and Sutton, J. 1982. Relaxing price competition through product differentiation. The Review of Economic Studies 49 (1), 3–13.

Shaked, A. and Sutton, J. 1983. Natural oligopolies. Econometrica 51 (5), 1469–1483.

Spence, M. 1975. Monopoly, Quality, and Regulation. Bell Journal of Economics 6 (2), 417–429.

Spence, M. 1976. Product selection, fixed costs, and monopolistic competition. The Review of Economic Studies 43 (2), 217–235.

Stenbacka, R. and Tombak, M. 2018. Optimal reimbursement policy in health care: Competition, ownership structure and quality provision. The B.E. Journal of Economic Analysis & Policy 18 (1), 1–19.

THL 2018. Terveydenhuollon menot ja rahoitus 2016. Statistical report. ⟨URL: http://urn.fi/URN:NBN:fi-fe2018070326792⟩.

United Nations 2017. National Accounts. ⟨URL: https://unstats.un.org/unsd/snaama/dnlList.asp⟩ [visited on 2018-10-13].

Urquiola, M. 2016. Competition among schools: Traditional public and private schools. In ed. by Hanushek, E. A., Machin, S., and Woessmann, L. Vol. 5. Handbook of the Economics of Education. Elsevier, 209–237.

Vandenbosch, M. B. and Weinberg, C. B. 1995. Product and price competition in a two-dimensional vertical differentiation model. Marketing Science 14 (2), 224–249.

Weizsacker, C. C. v. 1980. A welfare analysis of barriers to entry. The Bell Journal of Economics 11 (2), 399–420.

2 QUALITY AND COMPETITION BETWEEN PUBLIC AND PRIVATE FIRMS

Abstract[*]
We study a multistage, quality-then-price game between a public firm and a private firm. The market consists of a set of consumers who have different quality valuations. The public firm aims to maximize social surplus, whereas the private firm maximizes profit. In the first stage, both firms simultaneously choose qualities. In the second stage, both firms simultaneously choose prices. Consumers' quality valuations are drawn from a general distribution. Each firm's unit production cost is an increasing and convex function of quality. There are multiple equilibria. In some, the public firm chooses a low quality, and the private firm chooses a high quality. In others, the opposite is true. We characterize subgame-perfect equilibria. Equilibrium qualities are often inefficient, but under some conditions on con-sumer valuation distribution, equilibrium qualities are first best. Various policy implications are drawn.

Keywords: Price-quality competition, quality, public firm, private firm, mixed oligopoly.

[*] This chapter is joint work with Ching-to Albert Ma, Boston University. This chapter was published as *"Laine, L. and Ma, C.- t. A. 2017. Quality and competition between public and private firms. Journal of Economic Behavior & Organization 140, 336 – 353."* A version of this chapter has been published earlier as: *Laine, L. and Ma C.-t. A. 2016, Quality and competition between public and private firms,* in the Department of Economics Working Paper series. The paper is published here with a kind premission from Elsevier. For their suggestions and comments, we thank Francesca Barigozzi, Giacomo Calzolari, Helmuth Cremer, Vincenzo Denicolo, Katharina Huesmann, Ari Hyytinen, Qihong Liu, and the seminar participants at Auburn University, University of Bologna, Boston University, Helsinki Center of Economic Research, University of Hong Kong, National Taiwan University, the 16th European Health Economics Workshop in Toulouse, and the 14th International Industrial Organization Conference in Philadelphia. We also thank co-editor Tom Gresik, an associate editor, and two reviewers for their guidance. Laine is grateful for the Yrjö Jahnsson Foundation and the KAUTE Foundation for their financial support.

Journal of Economic Behavior & Organization 140 (2017) 336–353

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Quality and competition between public and private firms

Liisa T. Laine[1], Ching-to Albert Ma*

Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Received 27 July 2016
Received in revised form 9 May 2017
Accepted 11 May 2017
Available online 12 May 2017

JEL classification:
D4
L1
L2
L3

Keywords:
Price-quality competition
Quality
Public firm
Private firm
Mixed oligopoly

ABSTRACT

We study a multistage, quality-then-price game between a public firm and a private firm. The market consists of a set of consumers who have different quality valuations. The public firm aims to maximize social surplus, whereas the private firm maximizes profit. In the first stage, both firms simultaneously choose qualities. In the second stage, both firms simultaneously choose prices. Consumers' quality valuations are drawn from a general distribution. Each firm's unit production cost is an increasing and convex function of quality. There are multiple equilibria. In some, the public firm chooses a low quality, and the private firm chooses a high quality. In others, the opposite is true. We characterize subgame-perfect equilibria. Equilibrium qualities are often inefficient, but under some conditions on consumer valuation distribution, equilibrium qualities are first best. Various policy implications are drawn.

1. Introduction

Public and private firms compete in many markets. In many countries, general education, health care services and transportation are provided by public and private firms in various degrees. In higher education, most universities in Europe and Asia are public, but in the United States the market is a mixed oligopoly. Furthermore, in different markets in the U.S., quality segmentation varies. The best universities in the U.S. Northeast are private, but many public universities in California and the western states have higher quality than private colleges (see Deming and Goldin, 2012). In health care, again, many European markets are dominated by public firms, but in many countries the private market is very active. Again, quality segmentation differs. For example, in the U.S., according to the U.S. News ranking in 2016–2017, four out of the five best U.S. hospitals were private. However, it has been well documented that U.S. public nursing homes have higher quality than private nursing homes (see Comondore et al., 2009).

Quality is a major concern in these markets. The interest in quality stems from a fundamental point made by Spence (1975). Because a good's quality benefits all buyers, the social benefit of quality is the sum of consumers' valuations. At a social optimum, the average consumer quality valuation should be equal to the quality marginal cost. Yet, a profit- maximizing firm is only concerned with the consumer who is indifferent between buying and not. A firm's choice of quality will be one that maximizes the surplus of this marginal consumer. The classic Spence (1975) result says that even when products

* Corresponding author.
E-mail addresses: liisa.t.laine@jyu.fi (L.T. Laine), ma@bu.edu (C.-t.A. Ma).
[1] Permanent address: School of Business and Economics, P.O. Box 35, 40014 University of Jyväskylä, Finland.

http://dx.doi.org/10.1016/j.jebo.2017.05.012

are priced at marginal costs, their qualities will be inefficient. We show that a mixed oligopoly may be a mechanism for remedying this inefficiency.

We use a standard model of vertical product differentiation. In the first stage, two firms simultaneously choose product qualities. In the second stage, firms simultaneously choose product prices. Consumers' quality valuations are drawn from a general distribution. The two firms have access to the same technology. The only difference from the textbook setup is that one is a social-surplus maximizing public firm, whereas the other remains a profit- maximizing private firm. Surprisingly, this single difference has many implications.

First, the model exhibits multiple equilibria: in some equilibria, the public firm's product quality is higher than the private firm's, but in others, the opposite is true. These multiple equilibria illustrate the variety of quality segmentations in the markets mentioned above. Second, and more important, we present general conditions on consumers' quality-valuation distribution under which qualities in low-public-quality equilibria are efficient, as well as general conditions under which qualities in high-public-quality equilibria are efficient. When equilibrium qualities are inefficient, deviations from the first best go in tandem: either qualities in public and private firms are both below the corresponding first-best levels, or they are both above. Equilibrium qualities form a rich set, and we have constructed examples with many configurations.

Our analysis proceeds in the standard way. Given a subgame defined by a pair of qualities, we find the equilibrium prices. Then we solve for equilibrium qualities, letting firms anticipate that their quality choices lead to continuation equilibrium prices. In the pricing subgame, qualities are given. The public firm's objective is to maximize social surplus, so its price best response must achieve the efficient allocation of consumers across the two firms. This requires that consumers fully internalize the cost difference between high and low qualities. The public firm sets its price in order that the difference in prices is exactly the difference in quality costs. The private firm's best response is the typical inverse demand elasticity rule.

When firms choose qualities, they anticipate equilibrium prices in the next stage. Given the private firm's quality, the public firm chooses its quality to maximize social surplus, anticipating the equilibrium consumer assignment among firms in the next stage. The private firm, however, will try to manipulate the equilibrium prices through its quality. Without any price response from the public firm, the private firm would have chosen the quality that would be optimal for the marginal consumer, as in Spence (1975). A larger quality difference, however, would be preferred because that would raise the private firm's price. Because of the price manipulation, the private firm's equilibrium quality is one that maximizes the utility of an inframarginal consumer, not the utility of the marginal consumer.

In the first best, the socially efficient qualities are determined by equating average consumer valuations and marginal cost of quality. The surprise is that in contrast to private duopoly, the private firm's equilibrium quality choice may coincide with the first-best quality. In other words, the inframarginal consumer whose utility is being maximized by the private firm happens to have the average valuation among the private firm's customers.

The (sufficient) conditions for first-best equilibria refer to the consumers' quality-valuation distribution. In equilibria where the public firm produces at a low quality, equilibrium qualities are first best when the valuation distribution has a linear inverse hazard rate.[2] In equilibria where the public firm produces at a high quality, equilibrium qualities are first best when the valuation distribution has a linear inverse reverse hazard rate. The linear inverse hazard and inverse reverse hazard rate conditions are equivalent to the private firm's marginal revenue function being linear in consumer valuation. Nevertheless, linear inverse hazard and inverse reverse hazard rates are special. A generic valuation distribution violates linearity. In an inefficient equilibrium, both firms' qualities are either too high or too low relative to the first best. This is in sharp contrast to the private duopoly in which excessive quality differentiation is used to relax price competition.

We draw various policy implications from our results. First, if a public firm is to take over a private one in a duopoly, should it enter in the high-quality or low- quality segment of the market? High-public-quality equilibria and low-public-quality equilibria generate different social surpluses. Second, our use of a social-welfare objective function for the public firm can be regarded as making a normative point. If the public firm aims to maximize only consumer surplus, it will subscribe to marginal-cost pricing. Because the private firm never prices at marginal cost, equilibrium-price difference between firms will never be equal to the quality-cost difference, so consumer assignments across firms will never be efficient. A social-welfare objective does mean that the public firm tolerates high prices. However, our policy recommendation is that undesirable effects from high prices should be remedied by a tax credit or subsidy to consumers regardless of where they purchase from.

Our research contributes to the literature of mixed oligopolies. We use the classical model of quality-price competition in Gabszewicz and Thisse (1979, 1986) and Shaked and Sutton (1982, 1983). However, the mixed oligopoly literature revolves around the theme that a public firm may improve welfare. Grilo (1994) studies a mixed duopoly in the vertical differentiation framework. In her model, consumers' valuations of qualities follow a uniform distribution. The unit cost of production may be convex or concave in quality. The paper derives first-best equilibria. In a Hotelling, horizontal differentiation model with quadratic transportation cost, Cremer et al. (1991) show that a public firm improves welfare when the total number of firms is either two, or more than six. Also using a Hotelling model, Matsumura and Matsushima (2004) show that mixed oligopoly gives some cost-reduction incentives. In a Cournot model, Cremer et al. (1989) show the disciplinary effect of replacing some private firms by public enterprises. Comparing Cournot and Bertrand models in mixed market, Ghosh and Mitra (2010) show that the results from private Cournot-Bertrand comparisons do not hold when a private firm is replaced

[2] See Lemmas 3 and 6 below. If F denotes the distribution, and f the density, then the inverse hazard rate is $\frac{1-F}{f}$, and the inverse reverse hazard rate is $\frac{F}{f}$.

by a welfare- maximizing public firm. Our paper is consistent with these results. However, we use a general consumer valuation distribution and cost function, and present multiple equilibria, which have not been the focus in the literature.

For profit-maximizing firms, Cremer and Thisse (1991) show that, under very mild conditions on transportation costs, horizontal differentiation models are actually a special case of vertical product differentiation (see also Champsaur and Rochet, 1989). The isomorphism can be transferred to mixed duopolies. The key in the Cremer-Thisse (1991) proof is that demands in horizontal models can be translated into equivalent demands in vertical models. Firms' objectives are unimportant. Hence, results in horizontal mixed oligopolies do relate to vertical mixed oligopolies. In most horizontal differentiation models, consumers are assumed to be uniformly distributed on the product space, and the transportation or mismatch costs are quadratic. These assumptions translate to a uniform distribution of consumer quality valuations and a quadratic quality cost function in vertical differentiation models.

The first-best results in Grilo (1994) are related to the efficient equilibria in the two-firm case in Cremer et al. (1991) because both papers use the uniform distribution for consumer valuations. By contrast, we use a general distribution for consumer valuation. Our results simultaneously reveal the limitation of the uniform distribution and which properties of the uniform distribution (linear inverse hazard and linear inverse reverse hazard rates) have been the driver of earlier results. Furthermore, when consumer valuations follow a uniform distribution, the issue of multiple equilibria is moot for a duopoly. By contrast, we show that multiple equilibria are important for general distributions. Moreover, our equilibrium qualities translate to equilibrium locations under general consumer distributions on the Hotelling line.

For private firms, Anderson et al. (1997) give the first characterization for a general location distribution with quadratic transportation costs. Our techniques are consistent with those in Anderson et al. (1997), but we use a general cost function. A recent paper by Benassi et al. (2006) uses a symmetric trapezoid valuation distribution and explores consumers' nonpurchase options. Yurko (2011) works with lognormal distributions. Our monotone inverse hazard and inverse reverse hazard rate assumptions are valid under the trapezoid distribution, but invalid under lognormal distributions.

Qualities in mixed provisions are often discussed in the education and health sectors. However, perspectives such as political economy, taxation, and income redistribution are incorporated, so public firms typically are assumed to have objective functions different from social welfare. Brunello and Rocco (2008) combine consumers voting and quality choices by public and private schools, and let the public school be a Stackelberg leader. Epple and Romano (1998) consider vouchers and peer effects but use a competitive model for interaction between public and private schools. (For recent surveys on education and health care, see Urquiola, 2016; Barros and Siciliani, 2012.) Grassi and Ma (2011, 2012) present models of publicly rationed supply and private firm price responses under public commitment and noncommitment. Our results here indicate that commitment may not be necessary, and imperfectly competitive markets may sometimes yield efficient qualities.

Privatization has been a policy topic in mixed oligopolies. Ishibashi and Kaneko (2008) set up a mixed duopoly with price and quality competition. The model has both horizontal and vertical differentiation. However, all consumers have the same valuation on quality, and are uniformly distributed on the horizontal product space (as in Ma and Burgess, 1993). They show that the government should manipulate the objective of the public firm so that it maximizes a weighted sum of profit and social welfare, a form of partial privatization. (Using a Cournot model, Matsumura, 1998 earlier demonstrates that partial privatization is a valuable policy.) Our model is richer on the vertical dimension, but consists of no horizontal differentiation. Our policy implication has a privatization component to it, but a simple social welfare objective for the public firm is sufficient.

Section 2 presents the model. Section 3 studies equilibria in which the public firm's quality is lower than the private firm's, and Section 4 studies the opposite case. In each section, we first derive subgame-perfect equilibrium prices, and then equilibrium qualities. We present a characterization of equilibrium qualities, and conditions for equilibrium qualities to be first best. Section 5 considers policies, various robustness issues, and existence of equilibria. We consider alternative preferences for the public firm. We also let cost functions of the firms be different. Then we let consumers have outside options, and introduce multiple private firms. Finally we consider existence of equilibria. The last section presents some concluding remarks. Proofs are collected in the Appendix. Details of numerical computation are in the Supplement.

2. The model

2.1. Consumers

There is a set of consumers with total mass normalized at 1. Each consumer would like to receive one unit of a good or service. In our context, it is helpful to think of such goods and services as education, transportation, and health care including child care, medical, and nursing home services. The public sector often participates actively in these markets.

A good has a quality, denoted by q, which is assumed to be positive. Each consumer has a valuation of quality v. This valuation varies among consumers. We let v be a random variable defined on the positive support $[\underline{v}, \overline{v}]$ with distribution F and strictly positive density f. We also assume that f is continuously differentiable.

We will use two properties of the distribution, namely $[1-F]/f \equiv h$, and $F/f \equiv k$. We assume that h is decreasing, and that k is increasing, so $h'(v) < 0$ and $k'(v) > 0$. The assumptions ensure that profit functions, to be defined below, are quasi-concave, and are implied by f being logconcave (Anderson et al., 1997). These monotonicity assumptions are satisfied by many common distributions such as the uniform, the exponential, the beta, etc. (Bagnoli and Bergstrom, 2004). We call h the inverse hazard rate (because $1/h$ is the hazard), and k the inverse reverse hazard rate (because $1/k$ is the reverse hazard).

Valuation variations among consumers have the usual interpretation of preference diversity due to wealth, taste, or cultural differences. We may call a consumer with valuation v a type-v consumer, or simply consumer v. If a type-v consumer purchases a good with quality q at price p, his utility is $vq - p$. The quasi-linear utility function is commonly adopted in the literature (see, for example, the standard texts Anderson et al., 1992 and Tirole, 1988).

We assume that each consumer will buy a unit of the good. This can be made explicit by postulating that each good offers a sufficiently high benefit which is independent of v, or that the minimum valuation $\underline{v}$ is sufficiently high. The full-market coverage assumption is commonly used in the extant literature of product differentiation (either horizontal or vertical), but Delbono et al. (1996) and Benassi et al. (2016) have explored the implications of consumer outside options, and we defer to Subsection 5.4 for more discussions. Relatedly, the introduction of a public firm may be a policy for market expansion. We ignore this consideration by the full-market coverage assumption.

2.2. Public and private firms

There are two firms, Firm 1 and Firm 2, and they have the same technology. Production requires a fixed cost. The implicit assumption is that the fixed cost is so high that entries by many firms cannot be sustained. We focus on the case of a mixed oligopoly so we do not consider the rather trivial case of two public firms. Often a mixed oligopoly is motivated by a more efficient private sector, so in Subsection 5.3 we let firms have different technologies, and will explain how our results remain robust.

The variable, unit production cost of the good at quality q is $c(q)$, where $c : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing and strictly convex function. A higher quality requires a higher unit cost, which increases at an increasing rate. We also assume that c is twice differentiable, and that it satisfies the usual Inada conditions: $\lim_{q\to 0^+} c(q) = \lim_{q\to 0^+} c'(q) = 0$, so both firms always will be active.

Firm 1 is a public firm, and its objective is to maximize social surplus; the discussion of a general objective function for the public firm is deferred until Section 5.2. Firm 2 is a profit-maximizing private firm. Each firm chooses its product quality and price. We let p_1 and q_1 denote Firm 1's price and quality; similarly, p_2 and q_2 denote Firm 2's price and quality. Given these prices and qualities, each consumer buys from the firm that offers the higher utility. A consumer chooses a firm with a probability equal to a half if he is indifferent between them.

Consider any (p_1, q_1) and (p_2, q_2), and define $\hat{v}$ by $\hat{v}q_1 - p_1 = \hat{v}q_2 - p_2$. Consumer $\hat{v}$ is just indifferent between purchasing from Firm 1 and Firm 2. If $\hat{v} \in [\underline{v}, \bar{v}]$, then the demands for the two firms are as follows:

$$\begin{array}{ccl} \text{Demand for Firm 1} & \text{Demand for Firm 2} & \\ F(\hat{v}) & 1 - F(\hat{v}) & \text{if } q_1 < q_2 \\ 1 - F(\hat{v}) & F(\hat{v}) & \text{if } q_1 > q_2 \\ 1/2 & 1/2 & \text{if } q_1 = q_2 \end{array} \tag{1}$$

We sometimes call consumer $\hat{v}$ the indifferent or marginal consumer. (Otherwise, if $\hat{v} \notin [\underline{v}, \bar{v}]$, or fails to exist, one firm will be unable to sell to any consumer.)

If Firm 1's product quality is lower than Firm 2's, its demand is $F(\hat{v})$ when its price is sufficiently lower than Firm 2's price. Conversely, if Firm 2's price is not too high, then its demand is $1 - F(\hat{v})$. If the two firms' product qualities are identical, then they must charge the same price if both have positive demands. In this case, all consumers are indifferent between them, and each firm receives half of the market.

2.3. Allocation, social surplus, and first best

An allocation consists of a pair of product qualities, one at each firm, and an assignment of consumers across the firms. The social surplus from an allocation is

$$\int_{\underline{v}}^{v} [xq_\ell - c(q_\ell)]f(x)\mathrm{d}x + \int_{v}^{\bar{v}} [xq_h - c(q_h)]f(x)\mathrm{d}x. \tag{2}$$

Here, the qualities at the two firms are q_ℓ and q_h, $q_\ell < q_h$. Those consumers with valuations between $\underline{v}$ and v get the good with quality q_ℓ, whereas those with valuations between v and $\overline{v}$ get the good with quality q_h. The first best is (q_ℓ^*, q_h^*, v^*) that maximizes (2), and is characterized by the following:

$$\frac{\int_{\underline{v}}^{v^*} xf(x)\mathrm{d}x}{F(v^*)} = c'(q_\ell^*) \tag{3}$$

$$\frac{\int_{v^*}^{\overline{v}} xf(x)\mathrm{d}x}{1-F(v^*)} = c'(q_h^*) \tag{4}$$

$$v^*q_\ell^* - c(q_\ell^*) = v^*q_h^* - c(q_h^*). \tag{5}$$

In the characterization of the first best in (3), (4), and (5), those consumers with lower valuations should consume the good at a low quality (q_ℓ^*), and those with higher valuations should consume at a high quality (q_h^*). For the first best, divide consumers into two groups: those with $v \in [\underline{v}, v^*]$ and those with $v \in [v^*, \overline{v}]$. The (conditional) average valuation of consumers in $[\underline{v}, v^*]$ is in the left-hand side of (3), and, in the first best, this is equal to the marginal cost of the lower first-best quality, the right-hand side of (3). A similar interpretation applies to (4) for those consumers with higher valuations. Finally, the division of consumers into the two groups is achieved by identifying consumer v^* who enjoys the same surplus from both qualities, and this yields (5).

As Spence (1975) has shown, quality is like a public good, so the total social benefit is the aggregate consumer benefit, and in the first best, the average valuation should be equal to the marginal cost of quality. As a result the indifferent consumer v^* actually receives too little surplus from q_ℓ because $v^* > c'(q_\ell)$, but too much from q_h because $v^* < c'(q_h)$. In a private duopoly, firms will choose qualities to relax price competition, so one firm's equilibrium quality will be lower than the first best, whereas the other firm's equilibrium quality will be higher. Section 6 of Laine and Ma (2016), the working paper version, contains this result.[3]

2.4. Extensive form

We study subgame-perfect equilibria of the following game.

Stage 0: Nature draws consumers' valuations v and these are known to consumers only.
Stage 1: Firm 1 chooses a quality q_1; simultaneously, Firm 2 chooses a quality q_2.
Stage 2: Qualities in Stage 1 are common knowledge. Firm 1 chooses a price p_1; simultaneously, Firm 2 chooses a price p_2. Consumers then observe price-quality offers and pick a firm for purchase.

An outcome of this game consists of firms' prices and qualities, (p_1, q_1) and (p_2, q_2), and the allocations of consumers across the two firms. Subgames at Stage 2 are defined by the firms' quality pair (q_1, q_2). Subgame-perfect equilibrium prices in Stage 2 are those that are best responses in subgames defined by (q_1, q_2). Finally, equilibrium qualities in Stage 1 are those that are best responses given that prices are given by a subgame-perfect equilibrium in Stage 2.

There are multiple equilibria. In one class of equilibria, in Stage 1, the public firm chooses low quality, whereas the private firm chooses high quality, and in Stage 2, the public firm sets a low price, and the private firm chooses a high price. In the other class, the roles of the firms, in terms of their ranking of qualities and prices, are reversed. However, because the two firms have different objectives, equilibria in these two classes yield different allocations.[4]

3. Equilibria with low quality at public firm

3.1. Subgame-perfect equilibrium prices

Consider subgames in Stage 2 defined by (q_1, q_2) with $q_1 < q_2$. According to (1), each firm will have a positive demand only if $p_1 < p_2$, and there is $\tilde{v} \in [\underline{v}, \overline{v}]$ with

$$\tilde{v}q_1 - p_1 = \tilde{v}q_2 - p_2 \quad \text{or} \quad \tilde{v}(p_1, p_2; q_1.q_2) = \frac{p_2 - p_1}{q_2 - q_1}, \tag{6}$$

[3] Also, the multiple-quality duopoly with general valuation distributions and cost functions in Barigozzi and Ma (2016) can generate a special case for the single-quality duopoly with inefficient equilibrium qualities.

[4] These are all possible pure- strategy subgame-perfect equilibria. There is no equilibrium in which both firms choose the same quality. Indeed, the unique continuation equilibrium of subgames with identical qualities is firms setting price at the unit cost. Earning no profit, the private firm will deviate to another quality.

where we have emphasized that $\tilde{v}$, the consumer indifferent between buying from Firm 1 and Firm 2, depends on qualities and prices. Expression (6) characterizes firms' demand functions. Firm 1 and Firm 2s' payoffs are, respectively,

$$\int_{\underline{v}}^{\tilde{v}} [xq_1 - c(q_1)]f(x)\mathrm{d}x + \int_{\tilde{v}}^{\overline{v}} [xq_2 - c(q_2)]f(x)\mathrm{d}x \tag{7}$$

and

$$[1 - F(\tilde{v})][p_2 - c(q_2)]. \tag{8}$$

The expression in (7) is social surplus when consumers with valuations in $[\underline{v}, \tilde{v}]$ buy from Firm 1, whereas others buy from Firm 2. The prices that consumers pay to firms are transfers, so do not affect social surplus. The expression in (8) is Firm 2's profit.

Firm 1 chooses its price p_1 to maximize (7) given the demand (6) and price p_2. Firm 2 chooses price p_2 to maximize (8) given the demand (6) and price p_1. Equilibrium prices, $(\hat{p}_1, \hat{p}_2)$, are best responses against each other.

Lemma 1. *In subgames (q_1, q_2) with $q_1 < q_2$, and $\underline{v} < \frac{c(q_2)-c(q_1)}{q_2-q_1} < \overline{v}$, equilibrium prices $(\hat{p}_1, \hat{p}_2)$ are:*

$$\hat{p}_1 - c(q_1) = \hat{p}_2 - c(q_2) = (q_2 - q_1)\frac{1 - F(\hat{v})}{f(\hat{v})} \equiv (q_2 - q_1)h(\hat{v}), \tag{9}$$

$$\text{where } \hat{v} = \frac{c(q_2) - c(q_1)}{q_2 - q_1}. \tag{10}$$

In Lemma 1 the equilibrium price difference across firms is the same as the cost difference: $\hat{p}_2 - \hat{p}_1 = c(q_2) - c(q_1)$. Also, Firm 2 makes a profit, and its price-cost margin is proportional to the quality differential and the inverse hazard rate h.

We explain the result as follows. Firm 1's payoff is social surplus, so it seeks the consumer assignment to the two firms, $\tilde{v}$, to maximize social surplus (7). This is achieved by getting consumers to fully internalize the cost difference between the high and low qualities. Therefore, given $\hat{p}_2$, Firm 1 sets $\hat{p}_1$ so that the price differential $\hat{p}_2 - \hat{p}_1$ is equal to the cost differential $c(q_2) - c\,(q_1)$. In equilibrium, the indifferent consumer is given by $\hat{v}q_1 - c(q_1) = \hat{v}q_2 - c(q_2)$, which indicates an efficient allocation in the quality subgame (q_1, q_2).

Firm 2 seeks to maximize profit. Given Firm 1's price $\hat{p}_1$, Firm 2's optimal price follows the marginal-revenue-marginal-cost calculus. For a unit increase in p_2, the marginal loss is $[p_2 - c(q_2)]f(\tilde{v})/(q_2 - q_1)$, whereas the marginal gain is $[1 - F(\tilde{v})]$. Therefore, profit maximization yields $\hat{p}_2 - c(q_2) = (q_2 - q_1)\frac{1-F(\hat{v})}{f(\hat{v})}$, the inverse elasticity rule for Firm 2's price-cost margin.[5] Lemma 1 follows from these best responses.

The key point in Lemma 1 is that equilibrium market shares and prices can be determined separately. Once qualities are given, Firm 1 will aim for the socially efficient allocation, and it adjusts its price, given Firm 2's price, to achieve that. Firm 2, on the other hand, aims to maximize profit so its best response depends on Firm 1's price as well as the elasticity of demand. Firm 1 does make a profit, and we will return to this issue in Subsection 5.2.

To complete the characterization of price equilibria, we consider subgames (q_1, q_2) with $q_1 < q_2$, and either $\frac{c(q_2)-c(q_1)}{q_2-q_1} < \underline{v}$ or $\overline{v} < \frac{c(q_2)-c(q_1)}{q_2-q_1}$. In the former case, Firm 1 would like to allocate all consumers to Firm 2, whereas in the other case, Firm 1 would like to allocate all consumers to itself. In both cases, there are multiple equilibrium prices. They take the form of high values of $\hat{p}_1$ when all consumers go to Firm 2, but low values of $\hat{p}_1$ in the other. However, equilibria in the game must have two active firms, so these subgames cannot arise in equilibrium.

The equilibrium prices $(\hat{p}_1, \hat{p}_2)$ in (9) and (10) formally establish three functional relationships, those that relate any qualities to equilibrium prices and allocation of consumers across firms. We can write them as $\hat{p}_1(q_1, q_2)$, $\hat{p}_2(q_1, q_2)$, and $\hat{v}(q_1, q_2) \equiv \tilde{v}(\hat{p}_1(q_1, q_2), \hat{p}_2(q_1, q_2); q_1, q_2)$. We differentiate (9) with $\hat{v}$ in (10) to determine how equilibrium prices and market share change with qualities. As it turns out, we will only need to use the information of how $\hat{p}_1(q_1, q_2)$ and $\hat{p}_2(q_1, q_2)$ change with q_2.

Lemma 2. *From the definition of $(\hat{p}_1, \hat{p}_2)$ and $\hat{v}$ in (9) and (10), we have $\hat{v}$ increasing in q_1 and q_2, and*

$$\frac{\partial \hat{p}_1(q_1, q_2)}{\partial q_2} = h(\hat{v}) + h'(\hat{v})[c'(q_2) - \hat{v}] \tag{11}$$

$$\frac{\partial \hat{p}_2(q_1, q_2)}{\partial q_2} = c'(q_2) + h(\hat{v}) + h'(\hat{v})[c'(q_2) - \hat{v}]. \tag{12}$$

Lemma 2 describes how the equilibrium indifferent consumer changes with qualities, and the strategic effect of Firm 2's quality on Firm 1's price. The marginal consumer $\hat{v}$ is defined by $\hat{v}q_1 - c(q_1) = \hat{v}q_2 - c(q_2)$. Because $q_1 < q_2$, if q_1 increases, consumer $\hat{v}$ strictly prefers to buy from Firm 1, as does consumer $\hat{v} + \epsilon$ for a small and positive ϵ. Next, suppose that q_2

[5] Firm 2's demand is $1 - F(\tilde{v})$. Hence, elasticity is $\frac{\mathrm{d}(1-F(\tilde{v}))}{\mathrm{d}p_2}\frac{p_2}{1-F(\tilde{v})} = -\frac{q_2-q_1}{h(\tilde{v})}p_2$.

increases, consumer $\hat{v}$ also strictly prefers to buy from Firm 1. The point is that quality q_1 is too low for consumer $\hat{v}$ but quality q_2 is too high. An increase in q_1 makes Firm 1 more attractive to consumer $\hat{v}$, and an increase in q_2 makes Firm 2 less attractive to him.

If Firm 2 increases its quality, it expects to lose market share. However, it does not mean that its profit must decrease. From (8), Firm 2's profit is increasing in Firm 1's price.[6] If in fact Firm 1 raises its price against a higher q_2, Firm 2 may earn a higher profit. In any case, because h is decreasing, and $c'(q_2) > \hat{v}$, according to Lemma 2, an increase in q_2 may result in higher or lower equilibrium prices. The point is simply that Firm 2 can influence Firm 1's price response. Also, from the difference between (12) and (11), Firm 2's equilibrium price always increases at a higher rate than Firm 1's: $\partial\hat{p}_2/\partial q_2 - \partial\hat{p}_1/\partial q_2 = c'(q_2)$.

3.2. Subgame-perfect equilibrium qualities

At qualities q_1 and q_2, the continuation equilibrium payoffs for Firms 1 and 2 are, respectively,

$$\int_{\underline{v}}^{\hat{v}(q_1,q_2)} [xq_1 - c(q_1)]f(x)\mathrm{d}x + \int_{\hat{v}(q_1,q_2)}^{\bar{v}} [xq_2 - c(q_2)]f(x)\mathrm{d}x, \quad \text{and} \tag{13}$$

$$[1 - F(\hat{v}(q_1, q_2))][\hat{p}_2(q_1, q_2) - c(q_2)], \tag{14}$$

where $\hat{p}_2$ is Firm 2's equilibrium price and $\hat{v}$ is the indifferent consumer from Lemma 1. Let $(\hat{q}_1, \hat{q}_2)$ be the equilibrium qualities. They are mutual best responses, given continuation equilibrium prices:

$$\hat{q}_1 = \operatorname{argmax}_{q_1} \int_{\underline{v}}^{\hat{v}(q_1,\hat{q}_2)} [xq_1 - c(q_1)]f(x)\mathrm{d}x + \int_{\hat{v}(q_1,\hat{q}_2)}^{\bar{v}} [x\hat{q}_2 - c(\hat{q}_2)]f(x)\mathrm{d}x \tag{15}$$

$$\hat{q}_2 = \operatorname{argmax}_{q_2} [1 - F(\hat{v}(\hat{q}_1, q_2))][\hat{p}_2(\hat{q}_1, q_2) - c(q_2)]. \tag{16}$$

A change in quality q_1 has two effects on social surplus (13). First, it directly changes $vq_1 - c(q_1)$, the surplus of consumers who purchase the good at quality q_1. Second, it changes the equilibrium prices and the marginal consumer $\hat{v}$ (hence market shares) in Stage 2. This second effect is second order because the equilibrium prices in Stage 2 maximize social surplus. Hence, the first-order derivative of (13) with respect to q_1 is $\int_{\underline{v}}^{\hat{v}(q_1,q_2)}[x - c'(q_1)]f(x)\mathrm{d}x$ (although Firm 1's objective is to maximize social surplus of the entire market).

Similarly, a change in quality q_2 has two effects on Firm 2's profit. First, it directly changes the marginal consumer's surplus $\hat{v}q_2 - c(q_2)$. Second, it changes the equilibrium prices and the marginal consumer. We rewrite (16) as

$$[1 - F(\hat{v}(q_1, q_2))][\hat{v}(q_1, q_2)q_2 - c(q_2) - \hat{v}(q_1, q_2)q_1 + \hat{p}_1(q_1, q_2)] \tag{17}$$

because

$$\hat{v}(q_1, q_2) = \tilde{v}(\hat{p}_1(q_1, q_2), \hat{p}_2(q_1, q_2); q_1, q_2) \equiv \frac{\hat{p}_2(q_1, q_2) - \hat{p}_1(q_1, q_2)}{q_2 - q_1}, \tag{18}$$

which gives the channels for the influence of q_2 on prices. Firm 2's equilibrium price in Stage 2 maximizes profit, so the effect of q_2 on profit in (17) via $\hat{v}(q_1, q_2)$ has a second-order effect. Therefore, the first-order derivative of (17) with respect to quality q_2 is $\hat{v}(q_1, q_2) - c'(q_2) + \frac{\partial\hat{p}_1(q_1,q_2)}{\partial q_2}$ (where we have omitted the factor $[1 - F(\hat{v}(q_1, q_2))]$).

We set the first-order derivatives of social surplus with respect to q_1 and of profit with respect to q_2 to zero. Then we apply (11) in Lemma 2 to obtain the following.

Proposition 1. *Equilibrium qualities $(\hat{q}_1,\hat{q}_2)$, and the marginal consumer $\hat{v}$ solve the following three equations in q_1, q_2, and v*

$$\frac{\int_{\mathrm{v}}^{\mathrm{v}} xf(x)\mathrm{d}x}{F(v)} = c'(q_1)$$

$$v + \frac{h(v)}{1 - h'(v)} = c'(q_2)$$

$$vq_2 - c(q_2) = vq_1 - c(q_1).$$

Firm 1's objective is to maximize social surplus. However, given Firm 2's quality and the continuation equilibrium prices, the assignment of consumers across firms will always be efficient. Therefore, Firm 1's return to quality q_1 consists of the

[6] The partial derivative of (8) with respect to p_1 is $\frac{f(\tilde{v})[p_2 - c(q_2)]}{q_2 - q_1} > 0$.

benefits of its own consumers. Hence $\hat{q}_1$ equates the conditional average valuation of consumers in $[\underline{v}, \hat{v}]$, $\frac{\int_{\underline{v}}^{\hat{v}} xf(x)\mathrm{d}x}{F(\hat{v})}$, and the marginal cost $c'(q_1)$. This is the first equation.

Firm 2's quality will affect Firm 1's price in Stage 2. If this were not the case (imagine that $\partial\hat{p}_1/\partial q_2$ were 0), the profit-maximizing quality would be the optimal level for the marginal consumer: $\hat{v} = c'(q)$, reminiscent of the basic property of quality in Spence (1975). By raising quality from one satisfying $\hat{v} = c'(q)$, Firm 2 may also raise Firm 1's price, hence its own profit. This is a first-order gain. The optimal tradeoff is now given by $\hat{v} + \frac{\partial\hat{p}_1(\hat{q}_1,\hat{q}_2)}{\partial q_2} = c'(\hat{q}_2)$. We use (11) to simplify, and show that Firm 2 sets its quality to be efficient for a consumer with valuation $\hat{v} + \frac{h(\hat{v})}{1-h'(\hat{v})}$. This is the second equation.

Proposition 1 presents remarkably simple equilibrium characterizations. The only difference between equilibrium qualities and those in the first best stems from how Firm 2 chooses its quality. Firm 2's consumers have average valuation $\frac{\int_{\hat{v}}^{\bar{v}} xf(x)\mathrm{d}x}{1-F(\hat{v})}$, which should be set to the marginal cost of Firm 2's quality for social efficiency. However, Firm 2's profit-maximization objective leads it to set quality so that the marginal cost is equal to $\hat{v} + \frac{h(\hat{v})}{1-h'(\hat{v})}$. Our next result gives a class of valuation distributions for which the answer is affirmative. First, we present a mathematical lemma, which, through a simple application of integration by parts, allows us to write the conditional expectation of valuations in terms of inverse hazard rate and the density.

Lemma 3. *For any distribution F (and its corresponding density f and inverse hazard rate $h \equiv (1-F)/f$),*

$$\frac{\int_v^{\bar{v}} xf(x)\mathrm{d}x}{1-F(v)} \equiv v + \frac{\int_v^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)}. \tag{19}$$

Proposition 2. *Suppose that the inverse hazard rate h is linear; that is, $h(x) = \alpha - \beta x$, $x \in [\underline{v}, \bar{v}]$, for some α and $\beta \geq 0$. Then for any v*

$$v + \frac{h(v)}{1-h'(v)} = \frac{\int_v^{\bar{v}} xf(x)\mathrm{d}x}{1-F(v)} \equiv v + \frac{\int_v^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)}. \tag{20}$$

Equilibrium qualities and market shares are first best.

Proposition 2 exhibits a set of consumer valuation distributions for which the quality-price competition game yields first-best equilibrium qualities. We have managed to write the conditional average in terms of the inverse hazard rate in Lemma 3, and this is $\hat{v} + \frac{\int_{\hat{v}}^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(\hat{v})h(\hat{v})}$. When the inverse hazard rate is linear, $\frac{h(\hat{v})}{1-h'(\hat{v})} \equiv \frac{\int_{\hat{v}}^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(\hat{v})h(\hat{v})}$, Firm 2's profit-maximization incentive aligns with the social incentive. The following remark gives the economic interpretation for the linear inverse hazard rate.

Remark 1. When Firm 2 sells to high-valuation consumers, its marginal revenue is linear in consumer valuation if and only if $h(v)$ is linear.

The hazard rate has figured prominently in information economics and auction theory (see, for instance, Krishna, 2009; Laffont and Tirole, 1993; Myerson, 1997), and measures information rent, or virtual valuation. Here, in quality-price competition, its role is in how a private firm's quality changes the rival public firm's continuation equilibrium price. When the inverse hazard rate is linear, in auction and bargaining theory, strategies become linear and tractability is available (see Chatterjee and Samuelson, 1983; Gresik, 1991; Satterthwaite and Williams, 1989). Here, linear inverse hazard rate implies efficiency in equilibrium qualities.

We can use the differential equation $[1 - F(v)]/f(v) = \alpha - \beta v$ to solve for the valuation density.

Remark 2. Suppose that $h(x) = \alpha - \beta x$. Then if $\beta = 0$, f is the exponential distribution $f(x) = \frac{A}{\alpha}\exp(-\frac{x}{\alpha})$, with $\bar{v} = \infty$, and $A = \exp(\frac{\underline{v}}{\alpha})$, so when $\underline{v} = 0$, $f(x) = \frac{1}{\alpha}\exp(-\frac{x}{\alpha})$ for $x \in \mathbb{R}_+$. If $\beta > 0$, then $f(x) = \left[\frac{(\alpha-\beta x)^{(1-\beta)}}{(\alpha-\beta\underline{v})}\right]^{\frac{1}{\beta}}$, with $\alpha - \beta\bar{v} = 0$. For the uniform distribution, we have $h(x) = \bar{v} - x$ (so $\alpha = \bar{v}$, and $\beta = 1$).

Although equilibrium qualities are efficient when the inverse hazard rate is linear, Remark 2 shows that the set of valuation densities with linear inverse hazard rate is quite special—even among the set of two-parameter densities. The inverse hazard rate is unlikely to be linear for a randomly chosen distribution: the efficiency result in Proposition 2 may not be generic. What happens to qualities when they are inefficient? Our next result addresses that.

Proposition 3. *Let an equilibrium be written as $(\hat{q}_1, \hat{q}_2, \hat{v})$, corresponding to Firm 1's quality, Firm 2's quality, and the marginal consumer. If the equilibrium is not first best, either*

$$(\hat{q}_1, \hat{q}_2, \hat{v}) < (q_\ell^*, q_h^*, v^*) \quad \text{or} \quad (\hat{q}_1, \hat{q}_2, \hat{v}) > (q_\ell^*, q_h^*, v^*).$$

That is, when equilibrium qualities are not first best, either both firms have equilibrium qualities lower than the corresponding first-best levels, or both have equilibrium qualities correspondingly higher.

The proposition can be explained as follows. Firm 1 aims to maximize social surplus. If Firm 2 chooses $q_2 = q_h^*$, Firm 1's best response is to pick $q_1 = q_\ell^*$. Next, Firm 1's best response is increasing in q_2. This stems from the properties of $\hat{v}(q_1, q_2)$, the efficient allocation of consumers across the two firms. Quality q_1 is too low for consumer $\hat{v}$, whereas quality q_2 is too high. If q_2 increases, consumer $\hat{v}$ would become worse off buying from Firm 2, so actually $\hat{v}$ increases. This also means that Firm 1 should raise its quality because it now serves consumers with higher valuations. In other words, if Firm 2 raises its quality, Firm 1's best response is to raise quality. Therefore, Firm 1's quality is higher than the first best q_ℓ^* if and only if Firm 2's quality is higher than the first best q_h^*.

We can rewrite the equation for Firm 2's equilibrium quality choice as follows:

$$v + \left[\frac{\int_v^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)} - \frac{\int_v^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)} \right] + \frac{h(v)}{1 - h'(v)} = c'(q_2)$$

$$\left\{ \frac{h(v)}{1 - h'(v)} - \frac{\int_v^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)} \right\} + \frac{\int_v^{\bar{v}} x f(x)\mathrm{d}x}{1 - F(v)} = c'(q_2).$$

The term inside the curly brackets is the discrepancy in the characterization of the first-best high quality and Firm 2's equilibrium quality. We have provided a condition for this term to be zero in Proposition 2, but this cannot be expected to hold for most distributions. The property of this term will then determine the distortion described in Proposition 3.

We have constructed a number of examples to verify that equilibrium qualities can be either below or above the first best. However, it is more effective if we discuss these examples after we have presented the other class of equilibria in which the public firm chooses a higher quality than the private firm. The examples are presented in Section 4.3. Also, we will defer robustness and policy discussions until after we have presented the other class of equilibria, in Sections 5.1 and 5.2.

4. Equilibria with high quality at public firm

Because the two firms have different objectives, equilibria in this class are not isomorphic to those in the previous section. However, the logic of the analysis is similar to the previous subsections, so we will omit proofs (but some can be found in Laine and Ma, 2016).

4.1. Subgame-perfect equilibrium prices

When $q_1 > q_2$, the firms have positive demand only if $p_1 > p_2$. Now, consumers with high valuations buy from the public firm. We now write the definition of the indifferent consumer $\tilde{v}$ as:

$$\tilde{v}q_1 - p_1 = \tilde{v}q_2 - p_2 \quad \text{or} \quad \tilde{v}(p_1, p_2; q_1, q_2) = \frac{p_1 - p_2}{q_1 - q_2}. \tag{21}$$

Firm 1 and 2's payoffs are, respectively, social surplus and profit:

$$\int_{\underline{v}}^{\tilde{v}} [xq_2 - c(q_2)]f(x)\mathrm{d}x + \int_{\tilde{v}}^{\bar{v}} [xq_1 - c(q_1)]f(x)\mathrm{d}x \quad \text{and} \quad F(\tilde{v})[p_2 - c(q_2)].$$

Equilibrium prices, $(\hat{p}_1, \hat{p}_2)$, are best responses against each other, and characterized in the following lemma.

Lemma 4. *In subgames (q_1, q_2) with $q_1 > q_2$, and $\underline{v} < \frac{c(q_1)-c(q_2)}{q_1-q_2} < \bar{v}$, equilibrium prices $(\hat{p}_1, \hat{p}_2)$ are:*

$$\hat{p}_1 - c(q_1) = \hat{p}_2 - c(q_2) = (q_1 - q_2)\frac{F(\hat{v})}{f(\hat{v})} \equiv (q_1 - q_2)k(\hat{v}), \tag{22}$$

$$\text{where } \hat{v} = \frac{c(q_1) - c(q_2)}{q_1 - q_2}. \tag{23}$$

Firm 1 implements the socially efficient consumer allocation by setting a price differential equal to the cost differential, whereas Firm 2's profit maximization follows the usual marginal-revenue-marginal-cost tradeoff, which is now related to the inverse reverse hazard rate, $k = F/f$. Equilibrium prices, $\hat{p}_1(q_1, q_2)$, $\hat{p}_2(q_1, q_2)$ change with qualities in the following way.

Lemma 5. *From the definition of $(\hat{p}_1, \hat{p}_2)$ in (22) and (23), we have $\hat{v}$ increasing in q_1 and q_2,*

$$\frac{\partial \hat{p}_1(q_1, q_2)}{\partial q_2} = -k(\hat{v}) + k'(\hat{v})[\hat{v} - c'(q_2)], \tag{24}$$

$$\frac{\partial \hat{p}_2(q_1, q_2)}{\partial q_2} = c'(q_2) - k(\hat{v}) + k'(\hat{v})[\hat{v} - c'(q_2)]. \tag{25}$$

Unlike subgames where Firm 2's quality is higher than Firm 1's, Firm 2's market share increases with both q_1 and q_2. However, the effect of a higher quality q_2 on prices may be ambiguous, but the effect of q_2 on $\hat{p}_2$ is larger than that on $\hat{p}_1$ by $c'(q_2)$.

4.2. Subgame-perfect equilibrium qualities

Equilibrium qualities $(\hat{q}_1, \hat{q}_2)$ are mutual best responses, given continuation equilibrium prices:

$$\hat{q}_1 = \text{argmax}_{q_1} \int_{\underline{v}}^{\hat{v}(q_1,\hat{q}_2)} [x\hat{q}_2 - c(\hat{q}_2)]f(x)\text{d}x + \int_{\hat{v}(q_1,\hat{q}_2)}^{\overline{v}} [xq_1 - c(q_1)]f(x)\text{d}x \tag{26}$$

$$\hat{q}_2 = \text{argmax}_{q_2} F(\hat{v}(\hat{q}_1, q_2))[\hat{p}_2(\hat{q}_1, q_2) - c(q_2)], \tag{27}$$

where $\hat{p}_2$ is Firm 2's equilibrium price and $\hat{v}$ is the equilibrium indifferent consumer (see Lemma 4).

We apply the same method to characterize equilibrium qualities. Changing q_1 in Firm 1's payoff in (26) only affects the second integral there because the effect via the first integral is second order by the Envelope Theorem. Changing q_2 has only two effects: the direct effect on the surplus of the marginal consumer $\hat{v}q - c(q_2)$, and the effect on Firm 1's equilibrium price, because any effect on the marginal consumer is second order according to the Envelope Theorem. We obtain the first-order conditions

$$\int_{\hat{v}(q_1,q_2)}^{\overline{v}} [x - c'(q_1)]f(x)\text{d}x = 0 \text{ and } \hat{v}(\hat{q}_1, q_2) - c'(q_2) + \frac{\partial \hat{p}_1(\hat{q}_1, q_2)}{\partial q_2} = 0.$$

After applying Lemma 5 to the last first-order condition, we obtain the following.

Proposition 4. *Equilibrium qualities* $(\hat{q}_1,\hat{q}_2)$*, and the marginal consumer* $\hat{v}$ *solve the following three equations in* q_1, q_2, *and* v

$$\frac{\int_{v}^{\overline{v}} xf(x)\text{d}x}{1 - F(v)} = c'(q_1)$$

$$v - \frac{k(v)}{1 + k'(v)} = c'(q_2)$$

$$vq_2 - c(q_2) = vq_1 - c(q_1).$$

Proposition 4 shares the same intuition behind Proposition 1. Firm 1 chooses q_1 to maximize the surplus of those consumers with valuations higher than $\hat{v}$. Firm 2 chooses the quality that is efficient for a type lower than the marginal consumer, at valuation $\hat{v} - \frac{k(\hat{v})}{1+k'(\hat{v})}$. Firm 2's lower quality serves to use product differentiation to create a bigger cost differential, and hence a bigger price differential between the two firms.

We can identify a class of distributions for which Firm 2's profit incentive aligns with the social incentive. An intermediate result is the following.

Lemma 6. *For any distribution F (and its corresponding density f and inverse reverse hazard rate* $k \equiv F/f$*),*

$$\frac{\int_{\underline{v}}^{v} xf(x)\text{d}x}{F(v)} \equiv v - \frac{\int_{\underline{v}}^{v} f(x)k(x)\text{d}x}{f(v)k(v)}. \tag{28}$$

Proposition 5. *Suppose that the inverse reverse hazard rate k is linear; that is,* $k(x) = \gamma + \delta x$, $x \in [\underline{v}, \overline{v}]$*, for some* γ *and* $\delta \geq 0$*. Then for any* v

$$v - \frac{k(v)}{1 + k'(v)} = \frac{\int_{\underline{v}}^{v} xf(x)\text{d}x}{F(v)} \equiv v - \frac{\int_{\underline{v}}^{v} f(x)k(x)\text{d}x}{f(v)k(v)}. \tag{29}$$

Equilibrium qualities and market shares are first best.

The following two remarks, respectively, relate the linear inverse reverse hazard rate to the private firm's marginal revenue, and present the corresponding densities.

Remark 3. When Firm 2 sells to low-valuation consumers, its marginal revenue is linear in consumer valuation if and only if $k(v)$ is linear.

Remark 4. Suppose that $k(x) = \gamma + \delta x$. Then $\delta > 0$, and $f(x) = \left[\frac{(\gamma+\delta x)^{1-\delta}}{(\gamma+\delta \bar{v})}\right]^{\frac{1}{\delta}}$ with $\gamma + \delta \underline{v} = 0$. For the uniform distribution, $\gamma = -\underline{v}$ and $\delta = 1$.

Again, the above shows that densities that have linear reverse hazard rates constitute a small class. Generically, inefficient equilibrium qualities can be expected. When the equilibrium is not first best, the distortion in equilibria with higher public qualities exhibits the same pattern as in equilibria with lower public qualities: Proposition 3 holds verbatim for the class of high-public-quality equilibria: either both firms produce qualities higher than first best, or both produce qualities lower than first best.

4.3. Examples and comparisons between equilibrium and first-best qualities

Propositions 2, 3, and 5 point to a rich set of equilibrium qualities, which are often inefficient. Here, we construct a number of illustrative examples. We assume a quadratic cost function $c(q) = \frac{1}{2}q^2$. We consider six valuation distributions: two for each of triangular, truncated exponential, and beta distributions. For each distribution, we look at low-public-quality and high-public-quality equilibria. Diagrams 1–3 present the equilibrium qualities and social surpluses. (In each diagram, we mark the equilibrium and first-best qualities on a line, and write down the corresponding social surpluses to the right of the qualities.) Formulas of the inverse hazard and inverse reverse hazard rates and *Mathematica* programs are in the Supplement.

Example 1. A triangular distribution $f(v) = 2v$, and its reverse $f(v) = 2(1 - v)$, $v \in [0, 1]$.

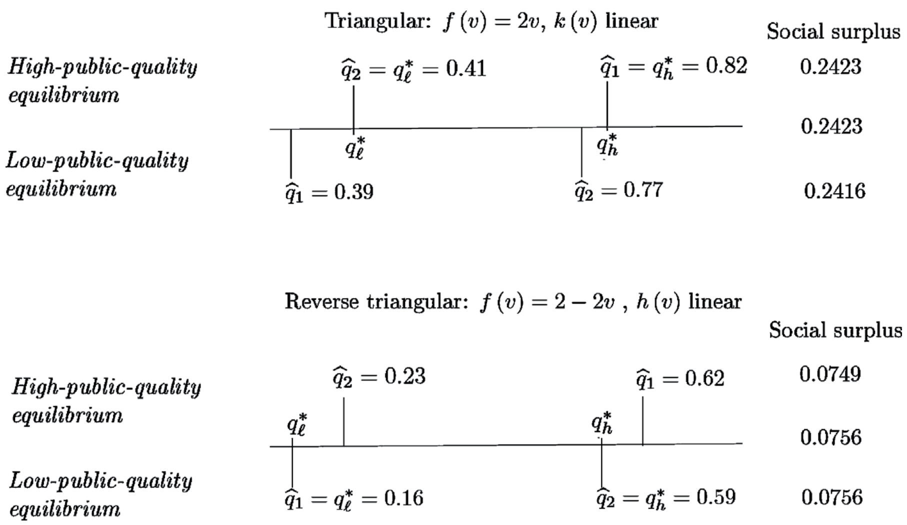

Diagram 1. Equilibria for triangular valuation distributions.

Example 1 shows the possibility of the first best. However, where equilibria are inefficient, qualities may be higher or lower than first best.

Example 2. A truncated exponential distribution $f(v) = \frac{[\exp(-v/\alpha)]/\alpha}{1-\exp(-\bar{v}/\alpha)}$, and its reverse $f(v) = \frac{[\exp(-(\bar{v}-v)/\alpha)]/\alpha}{1-\exp(-\bar{v}/\alpha)}$, $\alpha = 20$, and $v \in [0, \bar{v}] = [0, 100]$.

Example 2 shows that for the exponential distribution, equilibrium qualities must always be higher than first best, but for the reverse exponential distribution, equilibrium qualities must always be lower. The low-public-quality equilibrium yields a higher social surplus in the exponential distribution, but the reverse is true with the reverse exponential distribution.

Example 3. Two beta distributions: $f(v) = \frac{v^{(\alpha-1)}(1-v)^{(\beta-1)}}{\int_0^1 x^{(\alpha-1)}(1-x)^{(\beta-1)}dx}$, $v \in [0, 1]$, $(\alpha, \beta) = (2, 5)$ and $(\alpha, \beta) = (5, 2)$.

In Example 3, for each beta distribution, qualities are higher than first best in one equilibrium, but lower in the other. For the beta(2,5) distribution, the low- public-quality equilibrium yields a higher social surplus, but the revese is true for the beta(5,2) distribution.

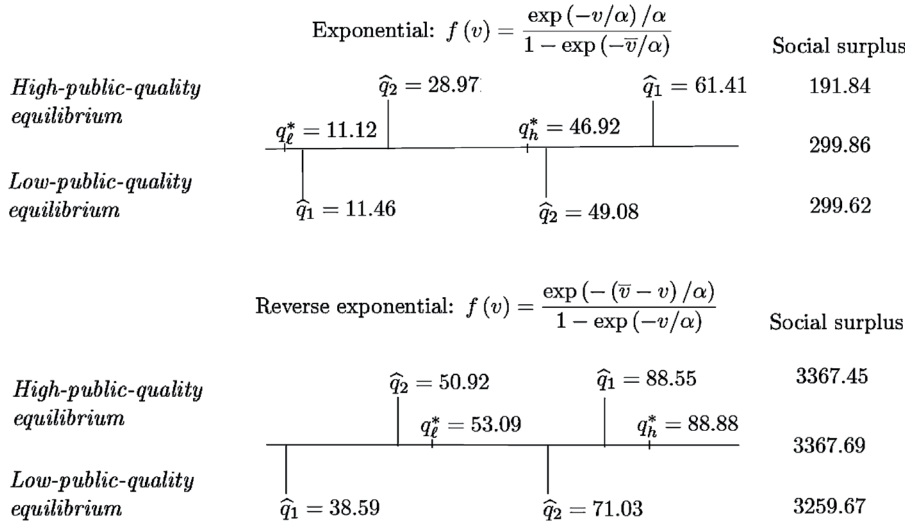

Diagram 2. Equilibria for truncated exponential valuation distributions.

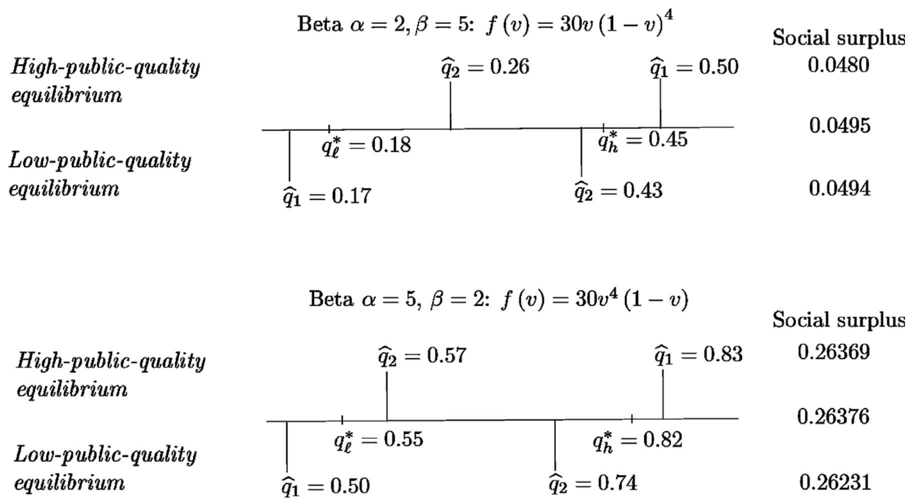

Diagram 3. Equilibria for beta valuation distributions.

5. Policies, robustness, and existence of equilibria

5.1. Competition policy

Suppose that the market initially consists of two private firms, so equilibrium qualities are inefficient. Qualities improve when one private firm is taken over by a public firm. Example 1 shows that for triangular and reverse triangular distributions, full efficiency can be restored if the public firm enters at the correct quality segment. Examples 2 and 3 show that generally low-public-quality and high-public-quality equilibria yield different social surpluses. Hence, entry by the public firm at the correct market segment is important. Our characterizations in Propositions 1 and 4 provide guidance.

Commitment by the public firm has been a common assumption in the previous literature. Equilibrium qualities are first best in the simultaneous-move games if and only if they are first best in the Stackelberg game, one in which the public firm can commit to quality or price. The reason is this. Suppose that Stackelberg equilibrium qualities are first best. Because the public firm's payoff is social surplus, the (first-best) low quality is a best response against the private firm's (first-best) high quality, so commitment is unnecessary. The converse is trivially true.

From Proposition 3, the improvement in welfare from a Stackelberg game comes from the public firm choosing a quality closer to the first best. For example, if in an equilibrium, qualities are lower than the first best (as in the reverse truncated exponential distribution case in Example 2), a higher public quality leads to a higher best response by the private firm, so both qualities will become closer to the first best.

5.2. General objective for the public firm and subsidies

So far our focus has been on quality efficiency. The public firm's objective function has been social welfare. Prices are transfers between consumers and firms, so do not affect social welfare. A more general objective function for a public firm can be a weighted sum of consumer surplus, and profits, also a common assumption in the literature. In this case, we can rewrite Firm 1's objective function as

$$\theta\left\{\int_{\underline{v}}^{\tilde{v}}[xq_1-p_1]f(x)\mathrm{d}x+\int_{\tilde{v}}^{\overline{v}}[xq_2-p_2]f(x)\mathrm{d}x\right\}+(1-\theta)\left\{F(\tilde{v})[p_1-c(q_1)]+[1-F(\tilde{v})][p_2-c(q_2)]\right\}. \quad (30)$$

Here, consumers are paying for the lower quality q_1 at price p_1, and the higher quality q_2 at price p_2. The weight on consumer surplus is $\theta>\frac{1}{2}$, whereas the weight on profits is $1-\theta$, so profits are unattractive from a social perspective. We can rewrite (30) as

$$\theta\left\{\int_{\underline{v}}^{\tilde{v}}[xq_1-c(q_1)]f(x)\mathrm{d}x+\int_{\tilde{v}}^{\overline{v}}[xq_2-c(q_2)]f(x)\mathrm{d}x\right\}-(2\theta-1)\left\{F(\tilde{v})[p_1-c(q_1)]+[1-F(\tilde{v})][p_2-c(q_2)]\right\},$$

which always decreases in Firm 1's price. If we impose a balanced-budget constraint, then the public firm must set price p_1 at marginal cost $c(q_1)$ to break even.

Lemmas 1 and 4 can no longer be valid. The first best cannot be an equilibrium because consumers do not bear the full incremental cost between high and low qualities. Suppose that $q_2>q_1$. The public firm will reduce price p_1 to marginal cost $c(q_1)$. However, in any price equilibrium, Firm 2's profit-maximizing price-cost margin has $p_2-c(q_2)>0$, so we have $p_2-p_1>c(q_2)-c(q_1)$. Fewer consumers will use the high-quality private firm.

We can regard the public firm's social-surplus preferences as a normative recommendation; otherwise, the distribution of consumers among firms will be inefficient. The concern for distribution should be addressed by a subsidy. Firms earn profits, according to Lemmas 1 and 4. Consider a low-public-quality equilibrium. Let equilibrium prices be $\hat{p}_1$ and $\hat{p}_2$. Impose taxes on Firms 1 and 2, respectively, at $F(\hat{v})[\hat{p}_1-c(q_1)]$ and $[1-F(\hat{v})][\hat{p}_2-c(q_2)]$, where $\hat{v}$ is in (10). The total tax revenue can be used as a consumer subsidy. For example, it can be equally distributed to all consumers, or be set up as a voucher for buying from either firm, or paid to consumers according to other criteria (say consumers with lower valuations get more). The only requirement is that the subsidy does not alter the difference of firms' prices, so that $p_2-p_1=c(q_2)-c(q_1)$, a necessary condition for the first best.

5.3. Different cost functions for public and private firms

We now let firms have different cost functions. Let $c_1(q)$ and $c_2(q)$ be Firm 1's and Firm 2's unit cost at product quality q, and these functions are increasing and convex.[7] The analysis in Sections 3 and 4 remains exactly the same. Simply replace every $c(q_1)$ by $c_1(q_1)$ and every $c(q_2)$ by $c_2(q_2)$. In the price subgame, the equilibrium still has a price difference equal to cost difference: $p_2-p_1=c_2(q_2)-c_1(q_1)$. The equilibrium qualities continue to satisfy their respective conditions after first-order conditions are simplified.

Propositions 2 and 5 have to be adjusted. This is because the first best in Section 2.3 has to be redefined. There are now two ways to assign technology. In one, low-valuation consumers pay the cost $c_1(q_\ell)$ for the low quality q_ℓ, and high-valuation consumers incur the cost $c_2(q_h)$ for the high quality q_h. In the other, it is the opposite. One of these technology assignments will yield a higher social welfare. However, our abstract model does not allow us to determine which technology should be used for low quality.[8]

The likelihood that the first best is achieved by the public firm taking over a private firm is small, again because linear inverse hazard and reverse hazard rates are nongeneric. The relevant question is whether the public firm should enter the low-quality segment or high-quality segement. Our examples for the case of identical cost functions show that the answer depends on the model specifics. This conclusion for competition policy should remain valid when costs are different.

5.4. Consumer outside option and many private firms

The consumer having an outside option is the same as introducing a fictitious firm offering a product at zero quality and zero price. In the first best, some consumers with very low valuations may not consume. The public firm's price affects

[7] Often the public firm is assumed to be less efficient. For example, we can let $c_1(q)>c_2(q)$ and $c_1'(q)>c_2'(q)$, so both unit and marginal unit costs are higher at the public firm. Our formal model, however, does not require this particular comparative advantage.

[8] As an illustration, let $c_1(q)=(1+s)c(q)$, and $c_2(q)=(1-s)c(q)$. The social welfare from using c_1 to produce the low quality is $\int_{\underline{v}}^{v}[xq_\ell-(1+s)c(q_\ell)]f(x)\mathrm{d}x+\int_{v}^{\overline{v}}[xq_h-(1-s)c(q_h)]f(x)\mathrm{d}x$. At $s=0$, this is the model in Subsection 2.1. From the Envelope Theorem, the derivative of the maximized welfare with respect to s, evaluated at $s=0$, is the partial derivative of welfare with respect to s: $-c(q_\ell^*)F(v^*)+c(q_h^*)[1-F(v^*)]$. Properties of q_ℓ^*, q_h^*, and v^* from (3), (4), and (5) do not indicate whether this derivative is positive or negative. It appears that the distribution F and the cost functions may interact in many ways.

decisions of two marginal consumers: the one who choose between the low-quality good and the high-quality good, and the one who choose between the low-quality good and no consumption at all.

In fact, Delbono et al. (1996) show that under a uniform valuation distribution, the first best is not an equilibrium. Efficient allocation requires that all consumers face price differentials equal to cost differentials. Hence, if Firm 1 produces a low quality q_1 and Firm 2 produces a high quality q_2, then efficiency requires $p_2 - p_1 = c(q_2) - c(q_1)$. When $p_2 > c(q_2)$ due to Firm 2's market power, $p_1 > c(q_1)$. However, to induce consumers to make efficient nonpurchase decisions, p_1 should be set at $c(q_1)$.

The case of many private firms is formally very similar. When a public firm has to interact with, say, two private firms, it does not have enough instruments to induce efficient decisions. Suppose that there are three firms, and that the medium quality is produced by a public firm, whereas the private firms produce low and high qualities. The public firm cannot simultaneously use one price to induce two efficient margins.

5.5. Existence of equilibria

In the previous sections, we have assumed the existence of equilibria. We now write down conditions for the solutions in Propositions 1 and 4 to be mutual best responses. For this, we consider two types of deviations: i) a firm choosing a lower quality than the rival's, and ii) a firm choosing a higher quality than the rival's.

Let $\pi_L(q_1) \equiv \max_{q_2 \le q_1} F(\hat{v}(q_1, q_2))[\hat{p}_2(q_1, q_2) - c(q_2)]$, where $\hat{v}(q_1, q_2) = [c(q_1) - c(q_2)]/[q_1 - q_2]$. Here, Firm 2 gets low-valuation consumers and the continuation equilibrium profit $\pi_L(q_1)$. Using the Envelope Theorem, we can show that $\pi_L(q_1)$ is strictly increasing.[9] If Firm 2 must choose only qualities that are lower than q_1, it benefits more when q_1 is higher because it has a bigger choice set. Let $\pi_U(q_1) \equiv \max_{q_2 \ge q_1} [1 - F(\hat{v}(q_1, q_2))][\hat{p}_2(q_1, q_2) - c(q_2)]$. Now, Firm 2 gets the high-valuation consumers and the profit $\pi_U(q_1)$. Again, using the Envelope Theorem, we can show that $\pi_U(q_1)$ is strictly decreasing. Firm 2's maximum profits from a continuation equilibrium is the upper envelope of $\pi_L(q_1)$ and $\pi_U(q_1)$, $\max\{\pi_L(q_1), \pi_U(q_1)\}$. Define $\overline{q}_1$ by $\pi_L(\overline{q}_1) = \pi_U(\overline{q}_1)$. The critical value $\overline{q}_1$ exists and is unique. It is a best response for Firm 2 to choose a high quality if and only if Firm 1's quality is below the critical value $\overline{q}_1$.

Next, for Firm 1's best response, we let $s_L(q_2) \equiv \max_{q_1 \le q_2} \int_{\underline{v}}^{\hat{v}(q_1,q_2)} [xq_1 - c(q_1)]f(x)\mathrm{d}x + \int_{\hat{v}(q_1,q_2)}^{\overline{v}} [xq_2 - c(q_2)]f(x)\mathrm{d}x$. This is the maximum social surplus when Firm 1's quality is lower than Firm 2's. Similarly, let $s_U(q_2) \equiv \max_{q_1 \ge q_2} \int_{\underline{v}}^{\hat{v}(q_1,q_2)} [xq_2 - c(q_2)]f(x)\mathrm{d}x + \int_{\hat{v}(q_1,q_2)}^{\overline{v}} [xq_1 - c(q_1)]f(x)\mathrm{d}x$, the maximum social surplus when Firm 1's quality is higher than Firm 2's. Again, using the Envelope Theorem, we show that $s_L(q_2)$ is strictly increasing, and $s_H(q_2)$ is strictly decreasing. Define $\overline{q}_2$ by $s_L(\overline{q}_2) = s_U(\overline{q}_2)$. It is a best response for Firm 1 to choose a low quality if and only if Firm 2's quality is above the critical value $\overline{q}_2$.

Formally, the low-public-quality equilibrium exists when the equations in Proposition 1 yield a solution $(\hat{q}_1, \hat{q}_2)$ satisfying $\hat{q}_1 < \overline{q}_1$ and $\hat{q}_2 > \overline{q}_2$. Similarly, the high-public-quality equilibrium exists when the equations in Proposition 4 yield a solution $(\hat{q}_1, \hat{q}_2)$ satisfying $\hat{q}_1 > \overline{q}_1$ and $\hat{q}_2 < \overline{q}_2$. However, we are unaware of general conditions on f and c for these requirements.

To confirm the existence of particular equilibria, however, we only need to verify that candidate equilibrium qualities are mutual best responses. For the $f(v) = 2v$ triangular distribution example above, we have computed each firm's payoffs. Given the private firm's quality q_2 set at a (candidate) equilibrium level, we compute the public firm's payoffs from setting quality q_1 at levels below and above q_2. We do the same for the private firm given the public firm's (candidate) equilibrium quality. We have confirmed, indeed, that those qualities in the example form an equilibrium. The computation details are in the Supplement. (We have also done the same for a model with v on a uniform distribution [10, 11] and a quadratic cost function. The game has an equilibrium with qualities at 10.25 and 10.75.)

6. Concluding remarks

In this paper we have studied equilibria in a mixed duopoly. The public firm maximizes social surplus, and the private firm maximizes profit. We have used a general distribution for consumer's valuations and a general cost function for firms. We discuss two classes of equilibria. In one class, the public firm offers low quality and the private offers high quality. In the other class, the opposite is true. Whereas generically, equilibrium qualities are inefficient, when inverse hazard or inverse reverse hazard rates are linear, equilibrium qualities are first best. We have related our results to competition policies, and discussed various robustness issues.

Various directions for further research may be of interest. Clearly, duopoly is a limitation. However, a mixed oligopoly with an arbitrary number of firms is analytically very difficult. In the extant literature, models of product differentiation with many private firms typically impose very strong assumptions on either consumer valuation (equivalently location) distribution or production cost (equivalently mismatch disutility). The contribution here relies on our ability to identify the inverse hazard and inverse reverse hazard rates as the determining factors for properties of equilibrium qualities. It may

---

[9] The derivative of $\pi_L(q_1)$ is the partial derivative of the profit function with respect to q_1 evaluated at the profit-maximizing q_2. This is $f(\hat{v})[\hat{p}_2 - c(q_2)]\frac{\partial \hat{v}}{\partial q_1} + F(\hat{v})\frac{\partial \hat{p}_2}{\partial q_1}$. We obtain $\frac{\partial \hat{p}_2}{\partial q_1}$ from (22) in Lemma 4. We verify that both $\frac{\partial \hat{v}}{\partial q_1}$ and $\frac{\partial \hat{p}_2}{\partial q_1}$ are positive, and conclude that $\pi_L(q_1)$ is strictly increasing. The monotonicity of $\pi_H(q_1)$, $s_L(q_2)$, and $s_H(q_2)$ can be demonstrated by similar computation.

well be that they also turn out to be useful for a richer model. The unit cost being constant with respect to quantity is a common assumption in the literature. We have used the same "constant-return" approach. Scale effects may turn out to be important even for the mixed duopoly.

Acknowledgements

For their suggestions and comments, we thank Francesca Barigozzi, Giacomo Calzolari, Helmuth Cremer, Vincenzo Denicolo, Katharina Huesmann, Ari Hyytinen, Qihong Liu, and the seminar participants at Auburn University, University of Bologna, Boston University, Helsinki Center of Economic Research, University of Hong Kong, National Taiwan University, the 16th European Health Economics Workshop in Toulouse, and the 14th International Industrial Organization Conference in Philadelphia. We also thank coeditor Tom Gresik, an associate editor, and two reviewers for their guidance. Laine received financial support from the Yrjö Jahnsson Foundation and the KAUTE Foundation.

Appendix

Proof of Lemma 1. Consider $\hat{p}_2 = \text{argmax}_{p_2}[1 - F(\tilde{v})][p_2 - c(q_2)]$, where $\tilde{v} = \frac{p_2 - \hat{p}_1}{q_2 - q_1}$ (see (6)). The first-order derivative of the profit function with respect to p_2 is

$$
\begin{aligned}
& [1 - F(\tilde{v})] - f(\tilde{v})[p_2 - c(q_2)]\frac{1}{q_2 - q_1} \\
= \quad & h(\tilde{v}) - [p_2 - c(q_2)]\frac{1}{q_2 - q_1},
\end{aligned}
$$

where we have used the partial derivative of $\tilde{v}$ with respect to p_2, namely $1/(q_2 - q_1)$. From the assumption that h is decreasing, the second-order derivative is negative, so the first-order condition is sufficient. Therefore, $\hat{p}_2$ is given by $\hat{p}_2 - c(q_2) = (q_2 - q_1)h(\tilde{v})$.

Next, consider Firm 1 choosing p_1 to maximize (7) where $\tilde{v} = \frac{\hat{p}_2 - p_1}{q_2 - q_1}$ (see (6)). Because (7) is independent of p_1, we can choose $\tilde{v}$ to maximize (7) ignoring (6). The optimal value $\hat{v}$ is given by setting to zero the first-order derivative of (7) with respect to $\tilde{v}$: $\hat{v}q_1 - c(q_1) = \hat{v}q_2 - c(q_2)$. Then we simply choose $\hat{p}_1$ to satisfy (6) such that $\hat{v} = \frac{\hat{p}_2 - \hat{p}_1}{q_2 - q_1} = \frac{c(q_2) - c(q_1)}{q_2 - q_1}$. We have shown that $\hat{p}_1$ and $\hat{p}_2$ in (9) and (10) are mutual best responses.

Proof of Lemma 2. First, from (10), we obtain $(q_2 - q_1)\text{d}\hat{v} + \hat{v}(\text{d}q_2 - \text{d}q_1) = c'(q_2)\text{d}q_2 - c'(q_1)\text{d}q_1$, which, together with the convexity of c, yields

$$\frac{\partial \hat{v}}{\partial q_1} = \frac{\hat{v} - c'(q_1)}{q_2 - q_1} = \frac{1}{q_2 - q_1}\left[\frac{c(q_2) - c(q_1)}{q_2 - q_1} - c'(q_1)\right] > 0 \tag{31}$$

$$\frac{\partial \hat{v}}{\partial q_2} = \frac{c'(q_2) - \hat{v}}{q_2 - q_1} = \frac{1}{q_2 - q_1}\left[c'(q_2) - \frac{c(q_2) - c(q_1)}{q_2 - q_1}\right] > 0. \tag{32}$$

Next, from (9), we obtain

$$
\begin{aligned}
\text{d}\hat{p}_1 - c'(q_1)\text{d}q_1 &= (\text{d}q_2 - \text{d}q_1)h(\hat{v}) + (q_2 - q_1)h'(\hat{v})\left(\frac{\partial \hat{v}}{\partial q_2}\text{d}q_2 - \frac{\partial \hat{v}}{\partial q_1}\text{d}q_1\right) \\
\text{d}\hat{p}_2 - c'(q_2)\text{d}q_2 &= (\text{d}q_2 - \text{d}q_1)h(\hat{v}) + (q_2 - q_1)h'(\hat{v})\left(\frac{\partial \hat{v}}{\partial q_2}\text{d}q_2 - \frac{\partial \hat{v}}{\partial q_1}\text{d}q_1\right).
\end{aligned}
$$

We then use (31) and (32) to simplify these, and obtain

$$
\begin{aligned}
\frac{\partial \hat{p}_1(q_1, q_2)}{\partial q_2} &= h(\hat{v}) + h'(\hat{v})\left[c'(q_2) - \hat{v}\right] \\
\frac{\partial \hat{p}_2(q_1, q_2)}{\partial q_2} &= c'(q_2) + h(\hat{v}) + h'(\hat{v})\left[c'(q_2) - \hat{v}\right],
\end{aligned}
$$

which are the expressions in the lemma.

Proof of Proposition 1. The first-order derivative of (13) with respect to q_1 is

$$\int_{\underline{v}}^{\hat{v}(q_1, q_2)} [x - c'(q_1)]f(x)\text{d}x + \left\{[\hat{v}(q_1, q_2)q_1 - c(q_1)] - [\hat{v}(q_1, q_2)q_2 - c(q_2)]\right\} f(\hat{v}(q_1, q_2))\frac{\partial \hat{v}}{\partial q_1}.$$

By Lemma 1, the term inside the curly brackets is zero. By putting this first-order derivative to zero, we obtain the first equation in the Proposition. Also, because equilibrium prices $\hat{p}_1(q_1, q_2)$ and $\hat{p}_2(q_1, q_2)$ must follow Lemma 1, we have

$$\hat{v}(q_1, q_2) = \frac{c(q_2) - c(q_1)}{q_2 - q_1},$$

which is the last equation in the Proposition.

Next, we use (17) to obtain the first-order derivative of Firm 2's profit with respect to q_2:

$$[1 - F(\hat{v}(q_1, q_2))] \left[\hat{v}(q_1, q_2) - c'(q_2) + \frac{\partial \hat{p}_1(q_1, q_2)}{\partial q_2}\right] + \left\{-f(\hat{v}(q_1, q_2))[\hat{p}_2(\hat{q}_1, q_2) - c(q_2)] + [1 - F(\hat{v}(q_1, q_2))](q_2 - q_1)\right\} \frac{\partial \hat{v}(q_1, q_2)}{\partial q_2}.$$

Again, by Lemma 1, the term inside the curly brackets is zero. After setting the first-order derivative to 0, we obtain

$$\hat{v}(q_1, q_2) - c'(q_2) + \frac{\partial \hat{p}_1(q_1, q_2)}{\partial q_2} = 0.$$

We then use (11) in Lemma 2 to substitute for $\frac{\partial \hat{p}_1(q_1, q_2)}{\partial q_2}$, and write the first- order condition as

$$\hat{v} - c'(q_2) + h(\hat{v}) + h'(\hat{v}) \left[c'(q_2) - \hat{v}\right] = 0,$$

which simplifies to

$$\hat{v} + \frac{h(\hat{v})}{1 - h'(\hat{v})} = c'(q_2),$$

the second equation in the Proposition.

Proof of Lemma 3. By definition, $f(x)h(x) = (1 - F(x))$. We have

$$\begin{aligned} & \frac{\int_v^{\bar{v}} xf(x)\mathrm{d}x}{1 - F(v)} \\ = & -\frac{\int_v^{\bar{v}} x\mathrm{d}(1 - F(x))}{f(v)h(v)} \\ = & \frac{v(1 - F(v))}{f(v)h(v)} + \frac{\int_v^{\bar{v}} (1 - F(x))\mathrm{d}x}{f(v)h(v)} \\ = & v + \frac{\int_v^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)}, \end{aligned}$$

where the second equality is due to integration by parts.

Proof of Proposition 2. Suppose that $h(x) = \alpha - \beta x$. We have, $h'(x) = - \beta$, and

$$v + \frac{h(v)}{1 - h'(v)} = v + \frac{\alpha - \beta v}{1 + \beta} = \frac{v + \alpha}{1 + \beta}.$$

Then we compute

$$v+\frac{\int_{v}^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)} = v+\frac{\int_{v}^{\bar{v}} f(x)(\alpha-\beta x)\mathrm{d}x}{f(v)h(v)}$$

$$= v+\frac{\alpha[1-F(v)]}{f(v)h(v)}-\beta\frac{\int_{v}^{\bar{v}} xf(x)\mathrm{d}x}{f(v)h(v)}$$

$$= v+\alpha-\beta\left\{v+\frac{\int_{v}^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)}\right\},$$

where the expression in the curly brackets comes from the identity (19). Simplifying, we have

$$v+\frac{\int_{v}^{\bar{v}} f(x)h(x)\mathrm{d}x}{f(v)h(v)}=\frac{v+\alpha}{1+\beta}.$$

We have proved (20).

The three equations in Proposition 1 are now exactly those that define the first best in (3), (4), and (5). Equilibrium qualities and consumer allocation must be first best.

Proof of Remark 1. When Firm 2 sells to consumers with valuations above v at price p_2, its revenue is $[1-F(v)]p_2$, where $v=\frac{p_2-p_1}{q_2-q_1}$. If we express p_2 as a function of v, we have $p_2(v)=p_1+v(q_2-q_1)$. The marginal revenue is the derivative of revenue with respect to the firm's quantity, $[1-F(v)]$:

$$\frac{\mathrm{d}[1-F(v)]p_2(v)}{\mathrm{d}[1-F(v)]}$$

$$= p_2(v)+[1-F(v)]\frac{\mathrm{d}p_2(v)}{\mathrm{d}[1-F(v)]}=p_2(v)+[1-F(v)]\frac{\mathrm{d}p_2(v)/\mathrm{d}v}{\mathrm{d}[1-F(v)]/\mathrm{d}v}$$

$$= p_2(v)-\frac{1-F(v)}{f(v)}\frac{\mathrm{d}p_2(v)}{\mathrm{d}v}=p_2(v)-h(v)(q_2-q_1).$$

Because $p_2(v)$ is linear in v, marginal revenue is linear in v if and only if the inverse hazard rate $h(v)$ is linear.

Proof of Remark 2. Define $y\equiv 1-F$, so $y'=-f$. We have $h(x)=\alpha-\beta x$ equivalent to $\frac{y'}{y}=\frac{-1}{\alpha-\beta x}$. First, suppose that $\beta=0$. We have $\frac{y'}{y}=\frac{-1}{\alpha}$, so $y(v)=A\exp(-\frac{v}{\alpha})$, some A. Therefore, $F(v)=1-A\exp(-\frac{v}{\alpha})$. Because we have $F(\underline{v})=0$, we must have $A=\exp(\frac{\underline{v}}{\alpha})$. We also have $F(\bar{v})=1$, which requires $\bar{v}=\infty$.

Second, suppose that $\beta>0$. We have $\frac{y'}{y}=\frac{-1}{\alpha-\beta v}$. Solving this differential equation, we have $y(v)=A(\alpha-\beta v)^{\frac{1}{\beta}}$, for some constant A. Hence, $F(v)=1-A(\alpha-\beta v)^{\frac{1}{\beta}}$, and we obtain the expression for f in the Remark by differentiation. Because $F(\underline{v})=0$, we have $A=(\alpha-\beta\underline{v})^{-\frac{1}{\beta}}$. Because $F(\bar{v})=1$, we must have $\alpha-\beta\bar{v}=0$, so that α and β cannot be arbitrary.

Proof of Proposition 3. For any q_2 we consider Firm 1's best response function:

$$\tilde{q}_1(q_2)=\mathrm{argmax}_{q_1}\int_{\underline{v}}^{\hat{v}(q_1,q_2)}[xq_1-c(q_1)]f(x)\mathrm{d}x+\int_{\hat{v}(q_1,q_2)}^{\bar{v}}[xq_2-c(q_2)]f(x)\mathrm{d}x.$$

First, at $q_2=q_h^*$, we have $\tilde{q}_1(q_h^*)=q_\ell^*$. Clearly, if Firm 2 chooses q_h^*, from the definition of the first best, Firm 1's best response is $q_1=q_\ell^*$ because Firm 1 aims to maximize social surplus. It follows that the first best belongs to the graph of Firm 1's best response function.

Second, we establish that $\tilde{q}_1(q_2)$ is increasing in q_2. The sign of the derivative of $\tilde{q}_1(q_2)$ has the same sign of the cross partial derivative of Firm 1's objective function (13) evaluated at $q_1=\tilde{q}_1(q_2)$. The derivative of (13) with respect to q_1 is simply

$$\int_{\underline{v}}^{\hat{v}(q_1,q_2)}[x-c'(q_1)]f(x)dx$$

because the partial derivative with respect to $\hat{v}$ is zero. The cross partial is then obtained by differentiating the above with respect to q_2, and this gives

$$[\hat{v}(q_1, q_2) - c'(q_1)]f(\hat{v})\frac{\partial \hat{v}(q_1, q_2)}{\partial q_2} > 0,$$

where the inequality follows because at $q_1 = \tilde{q}_1(q_2)$, we have $\hat{v}(q_1, q_2) > c'(q_1)$ and $\frac{\partial \hat{v}}{\partial q_2} > 0$ by (32) in the proof of Lemma 2.

Appendix A. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jebo.2017.05.012.

References

Anderson, S., de Palma, A., Thisse, J.-F., 1992. Discrete Choice Theory of Product Differentiation. MIT Press.
Anderson, S., Goeree, J., Ramer, R., 1997. Location, location, location. J. Econ. Theory 77, 102–127.
Bagnoli, M., Bergstrom, T., 2004. Log-concave probability and its applications. Econ. Theory 26, 445–469.
Barigozzi, F., Ma, C.-T.A., 2016. Product Differentiation with Multiple Qualities. Boston University Economics Department working paper.
Barros, P., Siciliani, L., 2012. Public and private sector interface. In: Pauly, M., McGuire, T., Barros, P. (Eds.), Handbook of Health Economics 2. , pp. 927–1002.
Benassi, C., Chirco, A., Colombo, C., 2006. Vertical differentiation and the distribution of income. Bull. Econ. Res. 58, 345–367.
Benassi, C., Castellani, M., Mussoni, M., 2016. Price equilibrium and willingness to pay in a vertically differentiated mixed duopoly. J. Econ. Behav. Organization 125, 86–96.
Brunello, G., Rocco, L., 2008. Educational standards in private and public schools. Econ. J. 118, 1866–1887.
Champsaur, P., Rochet, J.-C., 1989. Multiproduct duopolists. Econometrica 57, 533–557.
Chatterjee, K., Samuelson, W., 1983. Bargaining under incomplete information. Oper. Res. 31, 835–851.
Comondore, V., Devereaux, P., Zhou, Q., Stone, S., Busse, J., Ravindran, N., 2009. Quality of care in for-profit and not-for-profit nursing homes: systematic review and meta-analysis. Br. Med. J. 339.
Cremer, H., Thisse, J.-F., 1991. Location models of horizontal differentiation: a special case of vertical differentiation models. J. Ind. Econ. 39, 383–390.
Cremer, H., Marchand, M., Thisse, J.-F., 1989. The public firm as an instrument for regulating an oligopolistic market. Oxford Economic Papers New Series 41, 283–301.
Cremer, H., Marchand, M., Thisse, J.-F., 1991. Mixed oligopoly with differentiated products. Int. J. Ind. Organization 9, 43–53.
Delbono, F., Denicolò, V., Scarpa, C., 1996. Quality choice in a vertically differentiated mixed duopoly. Economic Notes by Banca Monte dei Paschi di Siena SpA 25, 33–46.
Deming, D., Goldin, C., 2012. The for-profit postsecondary school sector: nimble critters or agile predators? J. Econ. Perspect. 26, 139–164.
Epple, D., Romano, R., 1998. Competition between private and public schools, vouchers, and peer-group effects. Am. Econ. Rev. 88, 33–62.
Gabszewicz, J., Thisse, J.-F., 1979. Price competition, quality and income disparities. J. Econ. Theory 20, 340–359.
Gabszewicz, J., Thisse, J.-F., 1986. On the nature of competition with differentiated products. Econ. J. 96, 160–172.
Ghosh, A., Mitra, M., 2010. Comparing Bertrand and Cournot in mixed markets. Econ. Lett. 109, 72–74.
Grassi, S., Ma, C.-T.A., 2011. Optimal public rationing and price response. J. Health Econ. 30, 1197–1206.
Grassi, S., Ma, C.-T.A., 2012. Public sector rationing and private sector selection. J. Public Econ. Theory 14, 1–34.
Gresik, T., 1991. The efficiency of linear equilibria of sealed-bid double auctions. J. Econ. Theory 53, 173–184.
Grilo, I., 1994. Mixed duopoly under vertical differentiation. In: Annales d'Économie et de Statistique, No. 33, Concurrence, entreprises publiques et coopératives /Competition and Non- Profit-Maximizing., pp. 91–112.
Ishibashi, K., Kaneko, T., 2008. Partial privatization in mixed duopoly with price and quality competition. J. Econ. 95, 213–231.
Krishna, V., 2009. Auction Theory, 2nd edition. Academic Press.
Laffont, J.-J., Tirole, J., 1993. A Theory of Incentives in Procurement and Regulation. MIT Press.
Laine, L., Ma, C.-T.A., 2016. Quality and Competition between Public and Private Firms. Boston University Economics Department working paper, http://www.bu.edu/econ/files/2016/11/Ma_qual-comp_July2016.pdf.
Ma, C.-T.A., Burgess, J., 1993. Quality competition, welfare, and regulation. J. Econ. 58, 153–173.
Matsumura, T., 1998. Partial privatization in mixed duopoly. J. Public Econ. 70, 473–483.
Matsumura, T., Matsushima, N., 2004. Endogenous cost differentials between public and private enterprises: a mixed duopoly approach. Economica 71, 671–688.
Myerson, R., 1997. Game Theory Analysis of Conflict. Harvard University Press.
Satterthwaite, M., Williams, S., 1989. Bilateral trade with the sealed bid k-double auction: existence and efficiency. J. Econ. Theory 48, 107–133.
Shaked, A., Sutton, J., 1982. Relaxing price competition through product differentiation. Rev. Econ. Stud. 49, 3–13.
Shaked, A., Sutton, J., 1983. Natural oligopolies. Econometrica 51, 1469–1483.
Spence, A., 1975. Monopoly, quality and regulation. Bell J. Econ. 6, 417–429.
Tirole, J., 1988. The Theory of Industrial Organization. MIT Press.
Urquiola, M., 2016. Competition among schools: traditional public and private schools. In: Hanushek, E., Machin, S., Woessmann, L. (Eds.), Handbook of the Economics of Education Vol. 5. , pp. 209–237.
Yurko, A., 2011. How does income inequality affect market outcomes in vertically differentiated markets? Int. J. Industrial Organization 29, 493–503.

3 MULTI-DIMENSIONAL PRODUCT DIFFERENTIATION IN A MIXED OLIGOPOLY

Abstract*
I study a multistage two-attribute quality model of vertical product differentiation in a mixed duopoly: a social-surplus-maximizing public firm and a profit-maximizing private firm compete in two-attribute qualities and prices. I characterize the subgame-equilibrium prices and qualities of the model with general consumer valuation distributions and an increasing and convex per-unit cost function of quality. In contrast to the findings based on a single-attribute quality model, linearity in the inverse (or reverse) hazard of quality-valuation distribution is no longer a sufficient condition for establishing efficiency. Additional assumptions on the per-unit production cost of quality are required for the qualities to be socially optimal.

Keywords: Multi-dimensional product differentiation, price-quality competition, mixed oligopoly.

* I thank Anirban Basu, Ari Hyytinen, Fahad Khalil, Ching-to Albert Ma, and the participants of PHEnOM Seminar at the University of Washington for their comments. I undertook some of this research when I visited Columbia University; I am grateful for their generosity. I wish to thank Yrjö Jahnsson Foundation for financial support.

3.1 Introduction

When goods and services are provided in private oligopoly markets, the equilibrium quality does not coincide with the socially optimal quality. This result, originating from Spence (1975) and Moorthy (1988), has inspired a large literature on how public intervention would result in welfare improvements. A subset of this literature explores whether the presence of a social-surplus maximizing public firm improves the efficiency of market outcome. A key finding of this literature is that a public firm improves the market efficiency and that the equilibrium qualities may coincide with the socially optimal qualities (Cremer et al. 1991; Grilo 1994; Laine and Ma 2017).

The provision of goods simultaneously by public and private firms is common in many markets, for example in health care and education.[1] In these markets, product quality also often consists of multiple attributes. For example, quality of a non-urgent care or the long-term care provider can consists of clinical quality, amenities, waiting time, and geographic location of the care provider. In the case of education, students value academic reputation, teaching quality, general academic environment, facilities, and the location of the school differently. This paper contributes the existing literature by studying whether the presence of a social-surplus maximizing firm in the market improves the efficiency in the market when the quality of a good consists of multiple attributes.

When product quality has multiple attributes, an additional question to ask is whether firms in a mixed oligopoly differentiate in all dimensions in equilibrium. In private oligopolies, firms use (excess) product differentiation to relax price competition (Gabszewicz and Thisse 1979; Shaked and Sutton 1982, 1983). However, if a product quality consists of multiple attributes, private firms do not necessarily differentiate in all of the quality attributes (Barigozzi and Ma 2018). The competition effect can be different when firms have different objectives. In particular, the price of the public firm work as an instrument with a goal to induce or improve efficiency (Cremer et al. 1991; Grilo 1994; Laine and Ma 2017). Laine and Ma (2017) provide general (sufficient) conditions for this efficiency of the equilibrium qualities in a mixed duopoly: the equilibrium qualities are efficient if the consumers' quality-valuations have a linear inverse hazard or a linear reverse hazard rate. In the model studied by Laine and Ma (2017), however, the quality consists of one attribute. It turns out that in a two-attribute quality model, additional assumptions are needed to establish efficiency.

I build on a standard model of vertical product differentiation. In the first stage, firms choose their quality attributes. In the second stage, the qualities become common knowledge, and firms choose their prices. The model I use differs from a standard vertical differentiation model in two ways. First, a public firm maximizes social surplus and a private firm maximizes profits. In this part, my model follows closely the single-attribute quality mixed duopoly model by Laine

[1] Markets with a small number of firms and the objective function of at least one of the firms is different from others' are often called mixed oligopolies. (De Fraja and Delbono 1990)

and Ma (2017). Second, the quality of a product consists of two attributes. In this part, I follow closely the multi-attribute quality private duopoly model of vertical differentiation by Barigozzi and Ma (2018). These two extensions to the standard model of vertical product differentiation have many consequences.

Throughout the paper, I consider competition between a public and a private health care provider as an illustrative example. In this example, the first quality attribute may represent the quality of the amenities, and the second quality attribute may represent the (perceived) clinical quality.[2] In the context of this example, the first equilibrium class represents the equilibrium where a public firm provides a more modest quality on amenities than the private firm. In the second equilibrium class a public firm provides higher quality amenities than the private firm. I study whether in the equilibrium the health care providers provide different clinical qualities too and to which direction this potential differentiation in the equilibrium goes.

I proceed with the analysis as follows. I derive the equations that characterize the first best and the subgame equilibrium prices and qualities using a general distribution function for consumer valuations and an increasing and convex per-unit production cost of quality. To the best of my knowledge, my paper is the first to do so for a vertically differentiated two-attribute quality mixed duopoly. I characterize two classes of equilibria. In the first equilibrium class, the private firm provides higher quality in the first quality attribute, and the public firm provides lower quality in the first quality attribute. The second class of equilibrium considers the opposite. These two classes can be interpreted as different quality segments in the markets. Also, as shown by Laine and Ma (2017) for the single-attribute quality model, distinguishing between these two equilibrium classes is important because the equilibria may not be isomorphic, yielding different qualities and different social surplus.

After characterizing the first best and the two equilibrium classes, I add more structure to the model to obtain sharper results on product differentiation and efficiency of equilibrium qualities. I make two additional assumptions. First, I let the consumer valuations follow step functions. I also use a more specific form on the per-unit production cost of quality that allows me to study the effects of quality production cost spillovers on the equilibrium outcomes. Following the terminology used by Barigozzi and Ma (2018) I call the per-unit production cost of quality non-separable when there are positive or negative spillovers of quality production on costs. The focus on the per-unit cost of quality is motivated by Barigozzi and Ma (2018), who show that the separability in the per-unit production cost of quality plays a key role in the mixed differentiation results obtained from multi-attribute quality private duopoly models.

Using this version of the model, I show that if the per-unit cost of producing two quality attributes is non-separable in the two quality attributes, firms

[2] I assume that both quality dimensions are observable to the patients. Patients can use various tools for obtaining information about clinical quality measures such as health care quality report cards or clinical quality measures provided by The Centers for Medicare & Medicaid Services.

differentiate in both quality attributes. In particular, I give numerical examples using uniform quality-valuation distributions and a non-separable per-unit cost of quality and show that the equilibrium is not efficient. This proof by counterexample shows that unlike in the single-attribute quality mixed duopoly model, linearity in the inverse hazard or inverse reverse hazard cannot be the sufficient condition for efficiency. If this were the case, the equilibrium qualities would be efficient under uniform distributions. The numerical examples generate findings that contradict this statement. In all counterexamples the qualities are not efficient when the per-unit production costs of two quality attributes are not separable.

If the per-unit cost of producing two quality attributes is separable in the two quality attributes and the consumer valuations are uniform, the equilibrium qualities are the first best. This means that the equilibrium qualities being efficient in market with public and private providers suggested previously by Cremer et al. (1991), Grilo (1994), and Laine and Ma (2017) may also hold when quality consists of more than one quality attribute. The mechanism for this result is the following. When the quality production is separable, and the quality-valuations are uniform, the firms differentiate only in the first quality attribute. When the second attribute qualities are equal, the private firm's first-order condition with respect to the first quality attribute becomes exactly the same as it would be if there were only one quality attribute. The equations characterizing the equilibrium then simplify and become the same as in the single-attribute quality model presented in Laine and Ma (2017), and therefore their results regarding efficiency hold. I also give several numerical examples that confirm this result.

Last, I compare the inefficient equilibrium outcomes in the two classes of equilibria to the first best outcomes. My numerical examples show that when the equilibrium outcomes are inefficient, the equilibrium qualities and the social surplus can be different in the two equilibrium classes depending on whether the public offers high first quality attribute or the low first quality attribute. In case of the health care provider example with the first quality attribute representing the quality of the amenities and the second quality attribute representing the clinical quality, the results from my numerical examples show that the firms provide different qualities also in the clinical quality dimension in the equilibrium, and the equilibrium qualities are different depending on whether the public firm offered higher or lower amenity qualities than the private firm. These examples illustrate the property of the mixed oligopoly models, which highlights the importance of considering the different equilibrium classes separately.

These findings have implications for competition policy. Recall that the equilibrium qualities are generally inefficient in fully private markets (Moorthy 1988; Spence 1975). First, my results suggest that having a social-surplus maximizing firm in the market mean that in some cases equilibrium qualities in a two-attribute quality mixed duopoly can be efficient. A further question is that if instead of two private firms, one of the private firms in a private duopoly were replaced by a social surplus maximizing public firm, which quality segment of the market the public firm should choose? My results indicate that if the equilibrium qualities are not efficient even if the public firm is in the market, the presence of the public firm in the correct quality segment has important implications for social surplus: entering the "'incorrect"' (for example higher) quality segment may yield lower social surplus than if the public firm had chosen otherwise. Lastly, note that the differences between the two equilibrium classes arise even though the shape of the quality-valuation distribution is very symmetric (in my case uniform). The results of Laine and Ma (2017), however, suggest that other distributions for the quality-valuation with different shapes may affect the equilibrium outcome, and thus in a two-dimensional model they can even amplify the differences between two equilibrium classes.

I contribute to two strands of literature. The first strand of literature is the literature on multi-dimensional product differentiation in private duopolies in which differentiation is horizontal (Ansari et al. 1998; Irmen and Thisse 1998; Tabuchi 1994), vertical (Barigozzi and Ma 2018; Irmen and Thisse 1998; Lauga and Ofek 2011; Vandenbosch and Weinberg 1995), or both horizontal and vertical (Degryse and Irmen 2001).[3] A common finding in this literature is a max-min differentiation result, which means that in the two-attribute-quality models the firms differentiate in the first dimension and choose the same level of quality in the second dimension in equilibrium. Many of these assume uniform consumer valuations and a separable per-unit cost for quality production.

My notion of the importance of separability in the per-unit cost of quality is inspired by a recent article by Barigozzi and Ma (2018). They show that in a private duopoly with multi-attribute qualities firms differentiate in multiple dimensions (N-dimensional model) when the per-unit cost of quality is non-separable. The focus of Barigozzi and Ma (2018) is to establish a general differentiation result in a private duopoly. I focus instead on efficiency in markets with public and private providers.

The second strand of literature is the literature studying markets with public and private firms (mixed oligopolies). In a recent paper, Besley and Malcomson (2018) study entry by for-profit providers into market with unobserved quality and a non-profit incumbent.[4] I focus on price and quality competition and do not consider entry. In particular, my paper complements Laine and Ma (2017) by extending the model into a two-attribute quality model. The only paper studying two-attribute quality competition between firms with non-symmetric objectives is Rosenman and Munoz-Garcia (2017). They study competition between a market share maximizing firm and a profit-maximizing firm in a horizontally differentiated model but do not consider competition between a social-surplus maxi-

3 My paper is separate from product bundling literature (such as Chen and Riordan 2013). In my model, providers sell one good, and this good has several quality attributes, which cannot be sold separately. A bundled product, instead consists of several products that are sold as one combined product (Stole 2007, p.2281-2284).

4 See Laine and Ma (2017) for a more complete summary of the literature on public-private competition. Recently, Besley and Malcomson (2018) study entry by for-profit providers into market with unobserved quality and a non-profit incumbent. Stenbacka and Tombak (2018) study reimbursement policies and coverage in the health care markets with a profit-maximizing private firm and a non-profit firm.

mizing firm and a private firm. I use, instead, a more standard approach where the public firm maximizes social surplus. Moreover, in Rosenman and Munoz-Garcia (2017) the marginal unit production cost of quality is assumed to be zero and the consumer valuations are uniformly distributed. I focus on a model of vertical differentiation with less restrictive assumptions on the production costs and valuations. Thus, in contrast to Rosenman and Munoz-Garcia (2017) who find that firms choose not to differentiate in both quality dimensions, my results suggest that the firms in a mixed duopoly can offer products that are differentiated in both quality attributes.

I proceed as follows. Section 3.2 presents the model, and Section 3.3 characterizes the first best. Section 3.4 studies the equilibria in which the public firm's quality in the first quality attribute is lower than the private firm's quality attribute. Section 3.5 studies the opposite case. The last section concludes. Appendix 3.A provides proofs, derivations, and the description of the protocol I used in the numerical simulations. Mathematica programs used for numerical simulations are available from the author upon request.

3.2 The model

My modeling framework follows closely the multi-attribute quality private duopoly model of vertical differentiation by Barigozzi and Ma (2018) and the single-attribute quality mixed duopoly model by Laine and Ma (2017).

3.2.1 Consumers

There is a set of consumers, with the total mass normalized to one. A consumer buys one unit of a good or service. As an example, it is useful to think of education, transportation, and health care markets including child care, medical, and nursing home services, where public and private firms actively participate. Throughout the paper, I use competition between a public and a private health care provider as an illustrative example.

The quality of a good consists of a vector of two attributes $q \equiv (q_1, q_2) \in \mathbb{R}^2_+$.[5] In markets such as health care and education, it is common that a quality consists of multiple attributes. In my health care provider example, a quality of a provider for non-urgent care may consist of clinical quality in addition to other service related attributes, such as amenities and waiting time.

Consumers have heterogeneous preferences on quality attributes. Consumer preferences are represented by a vector of valuations $v \equiv (v_1, v_2) \in [\underline{v}_1, \overline{v}_1] \times [\underline{v}_2, \overline{v}_2] \in \mathbb{R}^2_+$. The valuation of q_1 is v_1 and the valuation of q_2 is v_2, and each valuation of quality attribute varies in a bounded and strictly positive interval.

5 With q, I refer to a general vector of product quality. At this point, I have not yet specified q to be provided by a particular firm. When analyzing the game, I use q to denote the quality offered by the public firm and r to denote the quality offered by the private firm.

The vector of valuations v is random, and each valuation v_i follows distribution function F_i with the corresponding density f_i, $i = 1, 2$. I sometimes use $\mathrm{d}F_i$ to denote the corresponding density of F_i. The valuation densities are continuously differentiable almost everywhere and log-concave. The assumption of log-concave densities implies that the joint density of $(v_1; v_2)$ is log-concave (Barigozzi and Ma 2018, footnote 3). Because of these assumptions Firm B's profits are quasiconcave in its own price (Caplin and Nalebuff 1991, Section 4). Last, I assume that the valuations on the two quality attributes are independent. This is a standard assumption in the literature of multi-attribute qualities (Barigozzi and Ma 2018; Degryse and Irmen 2001; Garella and Lambertini 2014; Lauga and Ofek 2011; Vandenbosch and Weinberg 1995).

The utility of a consumer with valuation vector v if he buys a good with a quality vector q with a price p is then $v_1q_1 + v_2q_2 - p$. The quasi-linear utility function is a standard assumption in the product differentiation literature (see for example Tirole 1988).

The market is covered. Assuming market coverage is a standard assumption in the product differentiation literature. The assumption of each consumer buying one good can be made explicit by either assuming that each good offers a sufficiently high valuation-independent benefit to the consumer, or that the minimum valuation of v_1 and v_2 is sufficiently high. Relaxing market coverage assumption would also complicate the model considerably as illustrated by Delbono et al. (1991) in the case of mixed oligopoly with uniform quality valuations and quadratic per-unit cost of quality.

3.2.2 Public and private firms

There are two firms in the market; Firm A and Firm B. There is a fixed cost for entering the market. This cost is high enough that it deters the entry of many private firms. Both firms have access to the same technology $C(q)$ to produce quality; the per-unit cost of quality $C : \mathbb{R}^2 \to \mathbb{R}_+$ is continuous, twice differentiable, strictly increasing, and a strictly convex function.

Firm A provides a good with the quality $q \equiv (q_1, q_2)$ and charges a price p_A. q_i refers to the level of Firm A's quality attribute i, $i = 1, 2$. Firm A maximizes social surplus: the sum of consumer′ surplus and profits (with equal weights on both). Social surplus maximization is a standard assumption in the mixed oligopoly literature (Cremer et al. 1991; Grilo 1994; Laine and Ma 2017).

Firm B provides a good with the quality $r \equiv (r_1, r_2)$ and charges a price p_B. r_i refers to the level of Firm B's quality attribute i. Firm B maximizes profits. When it sells a good with quality r at price p_B the per-unit profits are $p_B - C(r)$.

Given the quality choices of the firms, a consumer with valuation v who buys from Firm A receives utility $v_1q_1 + v_2q_2 - p_A$. Similarly, a consumer with valuation v who buys from Firm B receives utility $v_1r_1 + v_2r_2 - p_B$. A consumer buys from Firm A if and only if $v_1q_1 + v_2q_2 - p_A > v_1r_1 + v_2r_2 - p_B$. A consumer buys from Firm B if and only if $v_1q_1 + v_2q_2 - p_A < v_1r_1 + v_2r_2 - p_B$. The curve of consumers who are indifferent between buying from Firm A and Firm B is given by $v_1 q_1 + v_2 q_2 - p_A = v_1 r_1 + v_2 r_2 - p_B$. I solve this equation for v_1, and this defines the following function with $q_1 \neq r_1$:

$$\widetilde{v}_1 (v_2; p, q, r) = \frac{p_B - p_A}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}. \tag{1}$$

Function in (1) describes all consumers who are indifferent between buying from Firm A and B. Note that defining the curve of indifferent consumers by using v_1 is arbitrary. For example, the curve of indifferent consumers given by (1) could be solved by using v_2.

For the given vectors of qualities and prices the demands for Firm A and B are as follows

$$\begin{array}{ccc} \text{Demand for Firm } A & & \text{Demand for Firm } B \\ \int \int_{vq - p_A \geq vr - p_B} \mathrm{d}F_1 \mathrm{d}F_2 & \text{and} & \int \int_{vq - p_A \leq vr - p_B} \mathrm{d}F_1 \mathrm{d}F_2 \end{array}$$

Figure 1 depicts function (1) with an example in which both Firm B's quality attributes are higher than Firm A's quality attributes, that is $q_1 < r_1$, $q_2 < r_2$, and $p_A < p_B$. The curve of indifferent consumers divides the space of the two consumer valuation dimensions in two.

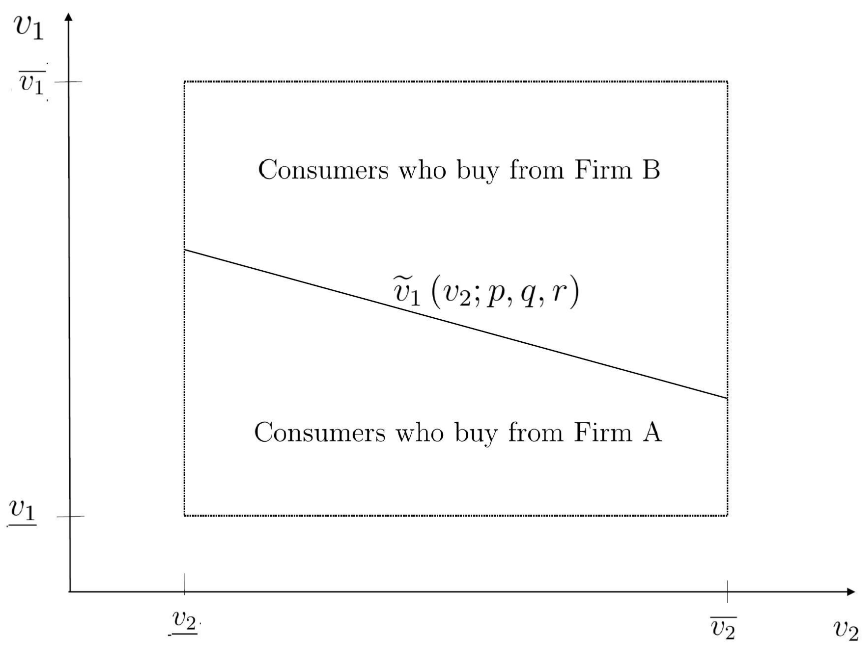

FIGURE 1 The curve of indifferent consumers $\widetilde{v}_1 (v_2; p, q, r)$ in a two-dimensional valuation space with $q_1 < r_1$, $q_2 < r_2$, and $p_A < p_B$ (Barigozzi and Ma, 2018).

Figure 1 illustrates the several features of the function $\widetilde{v}_1$ in (1). First, the curve of indifferent consumers $\widetilde{v}_1$ is linear in valuation v_2. This feature arises from the assumption of the quasi-linear consumer utility function. Second, definition in (1) shows that changes in prices affect only the intercept but not the slope of the curve of indifferent consumers. Keeping the qualities constant, the difference between prices p_A and p_B shifts the curve up or down. Third, qualities affect both intercept and the slope of $\widetilde{v}_1$.

Last, Figure 1 shows that the set of consumers with valuations below $\widetilde{v}_1$ buy from Firm A, that is a consumer (v_1', v_2) buys from Firm A if and only if $v_1' < \widetilde{v}_1 (v_2; p, q, r)$. In Figure 1, this is the set of consumers with valuations below $\widetilde{v}_1$. I call these consumers *lower first valuation consumers*. Similarly, a consumer (v_1', v_2) buys from Firm B if and only if $v_1' > \widetilde{v}_1 (v_2; p, q, r)$. In Figure 1, this is the set of consumers with valuations above $\widetilde{v}_1$. I call these consumers *higher first valuation consumers*. Following this logic, I can rewrite the demands above and collect them into the following:

$$\begin{array}{ll}
\text{Demand for Firm } A & \text{Demand for Firm } B \\
\text{If } q_1 < r_1 : & \\
\int_{\underline{v}_2}^{\overline{v}_2} \int_{\underline{v}_1}^{\widetilde{v}_1} \mathrm{d}F_1(v_1)\mathrm{d}F_2(v_2) & \int_{\underline{v}_2}^{\overline{v}_2} \int_{\widetilde{v}_1}^{\overline{v}_1} \mathrm{d}F_1(v_1)\mathrm{d}F_2(v_2) \\
\quad = \int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widetilde{v}_1\left(v_2; p, q, r\right)\right) \mathrm{d}F_2(v_2) & \quad = \int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widetilde{v}_1\left(v_2; p, q, r\right)\right)\right] \mathrm{d}F_2(v_2) \\
\text{If } q_1 > r_1 : & \\
\int_{\underline{v}_2}^{\overline{v}_2} \int_{\widetilde{v}_1}^{\overline{v}_1} \mathrm{d}F_1(v_1)\mathrm{d}F_2(v_2) & \int_{\underline{v}_2}^{\overline{v}_2} \int_{\underline{v}_1}^{\widetilde{v}_1} \mathrm{d}F_1(v_1)\mathrm{d}F_2(v_2) \\
= \int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widetilde{v}_1\left(v_2; p, q, r\right)\right)\right] \mathrm{d}F_2(v_2) & = \int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widetilde{v}_1\left(v_2; p, q, r\right)\right) \mathrm{d}F_2(v_2)
\end{array} \tag{2}$$

For $q = r$ and $p_A = p_B$, the firms sell to one half of the mass of consumers.

I avoid making unnecessary restrictions on how products can be differentiated. First, the products can be differentiated in the first quality attribute in two ways because both $q_1 < r_1$ and $q_1 > r_1$ are, a priori, possible. In the context of my health care provider example, it is possible that a public firm provides a more modest quality on amenities than the private firm and higher clinical quality, or vice versa. Second, it is possible that firms differentiate only in one quality attribute. For example, it is possible that a public firm provides lower quality on amenities than a private one, yet both provide equal clinical quality.

Referring back to Figure 1 and equation (1), when describing how products are differentiated relative to each other, I follow terminology from Garella and Lambertini (2014) and call quality attributes (or quality of a product) *r superior* if the quality is higher in both quality attributes, $q_1 < r_1$ and $q_2 < r_2$. When the quality is lower in both quality attributes, that is $q_1 < r_1$ and $q_2 < r_2$, I call quality *q inferior*. Firms can also differentiate so that the quality attributes are $q_1 < r_1$ and $r_2 < q_2$, or alternatively, $q_1 > r_1$ and $q_2 < r_2$. In these cases, I call quality attributes asymmetric.

The way products are differentiated relative to each other affects the slope of the curve of indifferent consumers. If $q_1 < r_1$, and if the goods of the two firms are such that one is strictly inferior and the other superior, the curve of indifferent consumers is downward sloping. When the quality attributes are asymmetric, the curve of indifferent consumers (1) is upward sloping. If there is differentiation only in one dimension, qualities in the second dimension are equal, and the slope of the set of indifferent consumers is constant.

3.2.3 Extensive form

I study the subgame-perfect equilibria of the following quality-price competition game:

Stage 0: Nature draws the consumer's valuations v from respective distributions. The valuations are only known to the consumer.

Stage 1: Firm A chooses the vector of product qualities q. Firm B chooses the vector of product qualities r.

Stage 2: Qualities chosen in Stage 1 become common knowledge. Firm A and Firm B simultaneously choose their prices. Firm A chooses p_A and Firm B chooses p_B. Consumers observe both firms' price-quality offers and pick one firm from which to buy.

I solve the game using backward induction. The outcome of the game of quality-price competition consists of firms' prices (p_A, p_B), vectors of qualities q and r, and an allocation of consumers across the two firms.

I characterize two classes of equilibria. In one class of equilibria, the public firm provides lower quality in the first quality attribute than the private firm, $q_1 < r_1$. In the second class of equilibria, the public firm provides higher quality in the first quality attribute than the private firm, $q_1 > r_1$.

It is important to distinguish between these two classes. In contrast to private duopoly in which both firms maximize profits, Firm A and Firm B now have different objectives; social surplus and profits respectively. Thus, equilibria in these two classes need not be isomorphic and may yield different outcomes. The importance of asymmetric objective functions for social surplus in a one dimensional mixed duopoly is also highlighted by Laine and Ma (2017), Sections 2.4 and 4.3.

There are no other classes of pure-strategy subgame-perfect equilibria than the two classes that I characterize. The reason is that if the public and private firms choose the equal qualities in both dimensions, the unique continuation equilibrium of subgames with an identical vector of qualities would be that each firm sets its the price at the per-unit (total) cost. This results from Firm A's objective to be social-surplus maximizing. However, if both firms set their prices at the per-unit cost of quality, the private firm would have an incentive to deviate to another quality to earn positive profit, and equal qualities would not be an equilibrium (see also Laine and Ma 2017, footnote 4 p).

In the analysis and characterizations below I focus on the necessary first-order conditions for social surplus and profit maximization. This means the characterizations are based on the assumption of the existence of a unique equilibrium in both subgames. Note that the game I consider here is a two-stage (finite) game of perfect information but with an infinite (continuous) actions in each stage. Showing generally existence is analytically demanding, as I am not aware of the non-trivial general conditions on the primitives that would ensure the existence of the subgame-perfect equilibria in pure strategies. Hellwig and Leininger (1987) show that the subgame-perfect equilibria in measurable pure strategies exists, when the payoff functions are continuous, strategy sets are compact, and constraint correnspondences are continuous. In case of the game studied in this paper, the existence of the subgame-perfect equilibrium could be established by showing that the conditions provided by Hellwig and Leininger (1987) hold. For a more detailed discussion of the issues related to existence, see Appendix 3.A.5.1.

I proceed under the assumption that nonexistence of an equilibrium is not an issue. For the numerical examples, I have checked that I have found the global maximum. A detailed description of the process is collected in Appendix 3.A.6.

3.3 Allocation, social surplus, and the first best

3.3.1 General definition of the first best

In this subsection I define the allocation, the social surplus, and characterize the first best with consumer valuation distributions, F_1 and F_2, and the per-unit production cost of quality C.

An allocation consists of two vectors of product qualities $\left(q^\ell, q^h\right)$, one vector each at Firm ℓ and Firm h, and an assignment of consumers across two firms, which will be determined by $\check{v}_1$. I assume that the social planner puts equal weight on consumers' surplus and firms' profit. I also assume that the second-order sufficient condition for the maximum which can be analyzed by using the successive principal minors of the symmetric Hessian determinant, such as I describe in Appendix A.5.1., holds.

The social surplus of an allocation is the sum of the consumer surplus and firms' profit:

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\check{v}_1} \left[v_1 q_1^\ell + v_2 q_2^\ell - C\left(q^\ell\right) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\check{v}_1}^{\overline{v}_1} \left[v_1 q_1^h + v_2 q_2^h - C\left(q^h\right) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2, \tag{3}$$

in which $q_1^\ell < q_1^h$. First best is $\left(q^{\ell *}, q^{h*}, v_1^*\right)$ that maximizes social surplus (3). The indexing of the quality attributes is arbitrary in the sense that the indices can be re-arranged without it changing the economic content of the first best.

The following equations (4)-(7), together with the function that determines the curve of indifferent consumers (8) characterize the first best:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{v_1^*(v_2;q^{\ell*},q^{h*})} v_1 \mathrm{d}F_1 \right\} \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(v_1^*\left(v_2;q^{\ell*},q^{h*}\right)\right) \mathrm{d}F_2} = C_1(q^{\ell*}) \tag{4}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{v_1^*(v_2;q^{\ell*},q^{h*})} v_2 \mathrm{d}F_1 \right\} \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(v_1^*\left(v_2;q^{\ell*},q^{h*}\right)\right) \mathrm{d}F_2} = C_2(q^{\ell*}) \tag{5}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{v_1^*(v_2;q^{\ell*},q^{h*})}^{\overline{v}_1} v_1 \mathrm{d}F_1 \right\} \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(v_1^*\left(v_2;q^{\ell*},q^{h*}\right)\right)\right] \mathrm{d}F_2} = C_1q^{h*}) \tag{6}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{v_1^*(v_2;q^{\ell*},q^{h*})}^{\overline{v}_1} v_2 \mathrm{d}F_1 \right\} \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(v_1^*\left(v_2;q^{\ell*},q^{h*}\right)\right)\right] \mathrm{d}F_2} = C_2(q^{h*}), \tag{7}$$

in which

$$v_1^*\left(v_2;q^{\ell*},q^{h*}\right) = \frac{C(q^{h*}) - C(q^{\ell*})}{q_1^{h*} - q_1^{\ell*}} - \frac{q_2^{h*} - q_2^{\ell*}}{q_1^{h*} - q_1^{\ell*}} v_2. \tag{8}$$

The first-best characterization divides consumers into two groups.[6] The division is achieved by identifying the set of consumers who enjoy equal surplus from both quality vectors. The set of consumers who enjoy equal surplus from both quality vectors is given by the function (8). The first group consists of consumers whose first dimension valuation is lower than the first dimension valuation of the indifferent consumers $v_1' < v_1^*\left(v_2;q^{\ell*},q^{h*}\right)$. The second group consists of consumers whose first dimension valuation is greater than the first dimension valuation of the indifferent consumers $v_1' > v_1^*\left(v_2;q^{\ell*},q^{h*}\right)$. A good's quality benefits all consumers who buy the good. Thus, the social benefit of a quality is the sum of the quality-valuations of those who buy the good. At a social optimum, the average consumer quality-valuation equals the marginal unit production cost of quality.

Equations (4) and (5) concern the lower first valuation consumers. The left-hand side of (4) is their (conditional) average of the first dimension quality-valuation. In the first best, this equals the first quality's ($q_1^{\ell*}$) marginal contribution to the per-unit production cost of the first dimension quality, the right-hand side of (4). The left-hand side of (5) is the (conditional) average valuation of the second dimension quality-valuation v_2 of the low first valuation consumers. In the first best, this equals the second quality's ($q_2^{\ell*}$) marginal contribution to the per-unit production cost of the second dimension quality, which is the right-hand side of (5).

Equations (6) and (7) concern the higher first valuation consumers. The left-hand side of (6) is their (conditional) average of the first dimension quality-valuation. In the first best this equals the first quality's (q_1^{h*}) marginal per-unit

6 I obtain (4)-(8) by solving for the first-order conditions with respect to q^ℓ, q^h, and $\check{v}_1$. For more details, see Appendix A.1.

contribution to unit production cost, the right-hand side of (6). The left-hand side of (7) is the (conditional) average valuation of the second dimension quality-valuation v_2. In the first best, this equals the second quality's (q_2^{h*}) per-unit marginal contribution to unit production cost, the right-hand side of (7).

In the first best, the average valuation of a quality equals the marginal contribution to the per-unit production cost of quality. Also, products are differentiated at least in one dimension. However, the characterization given by (4)-(8) does not say anything about whether it is efficient to differentiate in both quality dimensions. For example, in my health care provider example Firm ℓ may provide lower quality on amenities than Firm h, but based on (4)-(8) it is not clear whether providers differentiate in clinical qualities or not. The following subsection analyzes this in more detail.

3.3.2 Quality differentiation in the first best

My next result considers the first best when the per-unit production cost of quality has a more specific form: $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$. $C(q)$ exhibits positive cost spillovers from the production of two quality attributes if $\theta > 0$. $C(q)$ exhibits negative the cost spillovers from the production of two quality attributes if $\theta < 0$. $C(q)$ exhibits no cost spillovers from the production of two quality attributes if $\theta = 0$. When $\theta = 0$, the per-unit production cost function is called separable.

Result 1 *Suppose the per-unit production cost of quality is* $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$. *i) If* $\theta = 0$, *there is no differentiation in the second quality attribute, that is* $q_2^{\ell*} = q_2^{h*}$ *in the first best. ii) If* $\theta \neq 0$, *it is efficient to differentiate in both quality attributes, that is* $q_2^{\ell*} \neq q_2^{h*}$ *in the first best.*

Result 1 shows that for density functions differentiation in more than one quality attribute in the first best depends on the simplifying assumptions on the per-unit quality production function. If the per-unit production cost of quality is separable, it is efficient to differentiate only in one quality attribute. If the per-unit cost function of quality is not separable, efficiency requires differentiation in both quality attributes.

In my health care provider example, these results mean that if the per-unit production costs are separable, the firms differentiate in the amenity attribute but not in the clinical quality attribute in the first best. If the per-unit production costs are not separable, the firms differentiate in the amenity attribute and the clinical quality attribute in the first best.

3.3.3 Numerical examples of the first best when quality-valuation density is a step function

This subsection provides numerical examples that illustrate Result 1. I give an example in which there is differentiation only in one quality attribute and two examples in which qualities are differentiated in both quality attributes. I also use these numerical examples in Sections 3.4 and 3.5 to compare the equilibrium qualities to the efficient ones. I provide the details of the derivations of equations in this subsection in Appendix 3.A.4. For the details of the numerical examples, see Appendix 3.A.6.

For the numerical analysis that follows, I assume that consumers' valuations on two quality attributes, v_1 and v_2, are uniformly distributed on a strictly positive support $[1, 2]$. Also, the per-unit production cost of quality is quadratic: $C(q_1, q_2) = \frac{1}{2}q_1^2 + \theta q_1 q_2 + \frac{1}{2}q_2^2$.

Even though Result 1 applies for general density functions, I use uniform distribution in all my numerical examples as uniform distribution is the simplest valuation distribution with a linear inverse hazard (and reverse hazard rate). Uniform distribution is also a common assumption for quality-valuation in the product differentiation literature. I assume the support to be strictly positive because I want to make sure the qualities are positive (negative qualities are economically meaningless). With too narrow support and large θ values the solution yields corner solutions. Therefore, the support in the numerical examples is chosen as $[1, 2]$ instead of $[0, 1]$, for example. Too large θ values the solution yields corner solutions. Thus, θ values such neither of the equilibrium classes yield corner solutions. I have conducted simulations using several different values for θ and the results on differentiation and efficiently remain qualitatively the same.

When the quality-valuations are uniform and the per-unit cost production costs of quality is quadratic, (4)-(7) become:

$$\frac{1}{2}\left[\frac{1-\alpha^{*2}+3\alpha^*\beta^*-\frac{7}{3}\beta^{*2}}{1-\alpha^*+\frac{3}{2}\beta}\right] = q_1^{\ell*} + \theta q_2^{\ell*} \tag{9}$$

$$\frac{1}{2}\left[\frac{3-3\alpha^*+\frac{14}{3}\beta^*}{1-\alpha^*+\frac{3}{2}\beta^*}\right] = q_2^{\ell*} + \theta q_1^{\ell*} \tag{10}$$

$$\frac{1}{2}\left[\frac{2-\alpha^{*2}+3\alpha^*\beta^*-\frac{7}{3}\beta^{*2}}{2-\alpha^*+\frac{3}{2}\beta^*}\right] = q_1^{h*} + \theta q_2^{h*} \tag{11}$$

$$\frac{1}{2}\left[\frac{3-3\alpha^*+\frac{14}{3}\beta^*}{2-\alpha^*+\frac{3}{2}\beta^*}\right] = q_2^{h*} + \theta q_1^{h*}, \tag{12}$$

in which $\alpha^* \equiv \frac{\frac{1}{2}q_1^{h*2}+\theta q_1^{h*}q_2^{h*}+\frac{1}{2}q_2^{h*2}-\left[\frac{1}{2}q_1^{\ell*2}+\theta q\ell_1^* q_2^{\ell*}+\frac{1}{2}q^{\ell}{*_2^2}\right]}{q_1^{h*}-q_1^{\ell*}}$, and $\beta^* \equiv \frac{q_2^{h*}-q_2^{\ell*}}{q_1^{h*}-q_1^{\ell*}}$, and to clarify the notation I have omitted $(q^{\ell*}, q^{h*})$ from $\alpha(q^{\ell*}, q^{h*})$ and $\beta^*(q^{\ell*}, q^{h*})$ from the formulations.

Table 1 gives three different examples of the first best qualities and the social surplus for different parameter values of θ. The main findings from these examples are summarized in Remark 1.

TABLE 1 First best qualities and social surplus for $q_1^{\ell *} < q_1^h$.

Example	First best	qualities			Social surplus
(1) $\theta = 0$:	$q_1^{\ell *} = 1.250$	$q_2^{\ell *} = 1.500$	$q_1^{h*} = 1.750$	$q_2^{h*} = 1.500$	2.2812500
(2) $\theta = 0.005$:	$q_1^{\ell *} = 1.243$	$q_2^{\ell *} = 1.496$	$q_1^{h*} = 1.743$	$q_2^{h*} = 1.489$	2.2700583
(3) $\theta = -0.005$:	$q_1^{\ell *} = 1.258$	$q_2^{\ell *} = 1.504$	$q_1^{h*} = 1.758$	$q_2^{h*} = 1.511$	2.2925589

Notes. This table reports results for three different numerical examples. The quality-valuation distribution is uniform on $[1, 2]$ and the per-unit cost function of quality is quadratic. The three examples are distinguished by the value assigned to parameter θ, which is the cost spillover in producing quality attributes. The qualities refer to the first best qualities which maximize the social surplus. The social surplus refers to the maximum social surplus that the first best qualities yield.

Remark 1 *The numerical examples of Table 1 show that i) if $\theta = 0$, it is efficient to differentiate only in one dimension, and that ii) if $\theta \neq 0$, it is efficient to differentiate in both dimensions.*
In addition to confirming the findings from Result 1, the numerical examples of Table 1 show that iii) social surplus is the highest when there are negative spillovers of quality production on costs.

In Example 1 the per-unit production cost is separable. It confirms the non-differentiation result given by Result 1: when there are no cost spillovers from the production of two quality attributes, it is efficient to differentiate only in the first quality attribute. In this example, the efficient second attribute qualities are both 1.50, which is the average valuation of all the v_2 consumers (the midpoint of the whole support of v_2 that is $[1, 2]$). The efficient first attribute quality for the consumers that Firm ℓ serves is the average valuation of the support of v_1, $[1, 1.5]$.[7] This efficient lower first attribute quality is $q_1^{\ell *} = 1.25$. The efficient first attribute qualities for the consumers that Firm h serves is the average valuation of the support of v_1, $[1.5, 2]$. This efficient higher first attribute quality is $q_1^{h*} = 1.75$.

In Examples 2 and 3 there are non-zero cost spillovers from the production of two quality attributes. These examples confirm the differentiation result given by Result 1: if there are cost spillovers from the production of the two quality attributes ($\theta \neq 0$) there is differentiation in both quality attributes. In Example 2, the cost spillover from the production of two quality attributes is positive ($\theta > 0$). In Example 3, the cost spillover from the production of two quality attributes is negative ($\theta < 0$). Efficient qualities are different in each example. Moreover, in these examples the differentiation in the first best is purely asymmetric: in all examples Firm h produces a superior first quality compared to that of Firm ℓ, and Firm ℓ produces a superior second quality compared to that of Firm h.

Last, the numerical examples complements the results from Remark 1 by showing that the social surplus in the first best can vary depending on the values of θ. Moreover, in the examples above the social surplus is highest when the cost spillover from the production of two quality attributes is negative ($\theta = -0.005$).

7 In this example with the chosen parameter values, the first best $v^* = 1.5$.

3.4 Equilibria with low first quality attribute at the public firm

In this section, I characterize the class of equilibria where Firm A offers a lower first quality attribute than Firm B, that is $q_1 < r_1$. In my health care provider example this equilibrium class represents a class in which Firm A offers more modest quality amenities than Firm B. Note, that in the equilibrium it may be the case that differentiation happened in the amenities, and both firms would provide equal qualities. Also, differentiation in equilibrium may happen in a different dimension than what would be required for the first best. I characterize the second class of equilibria with $q_1 > r_1$ in Section 3.5.

In the following two subsections, I solve the subgame-perfect equilibrium prices and then continue by solving the subgame-perfect equilibrium qualities.

3.4.1 Subgame-perfect equilibrium prices

Suppose $q_1 < r_1$ and consider subgame (q, r) in Stage 2. Function

$$\widetilde{v}_1 \left(v_2; p_A, p_B, q, r\right) = \frac{p_B - p_A}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1} \tag{13}$$

defines all consumers who are indifferent between buying from Firm A and Firm B. This function depends on the prices charged by the firms and the quality attributes. Firm A's payoffs are

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\widetilde{v}_1(v_2;p,q,r)} \left[v_1 q_1 + v_2 q_2 - C\left(q\right)\right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\widetilde{v}_1(v_2;p,q,r)}^{\overline{v}_1} \left[v_1 r_1 + v_2 r_2 - C\left(r\right)\right] \mathrm{d}F_1 \right\} \mathrm{d}F_2, \tag{14}$$

with the vector of prices (p_A, p_B), and $\widetilde{v}_1$ is defined by (13). In the expression (14) consumers with valuations below $\widetilde{v}_1$ buy from Firm A. Consumers with valuations above $\widetilde{v}_1$ buy from Firm B.

Firm B's payoffs are

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\widetilde{v}_1(v_2;p,q,r)}^{\overline{v}_1} \left[p_B - C\left(r\right)\right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 = \int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widetilde{v}_1\left(v_2; p, q, r\right)\right)\right] \left[p_B - C\left(r\right)\right] \mathrm{d}F_2, \tag{15}$$

in which p denotes the vector of prices (p_A, p_B) and $\widetilde{v}_1$ is defined by (13).

Firm A chooses its price p_A to maximize (14) given the curve of indifferent consumers (13) and price p_B. Firm B chooses its price p_B to maximize (15) given the curve of indifferent consumers (13) and price p_A. Equilibrium prices $(\widehat{p}_A, \widehat{p}_B)$ are the best responses against one another.

Lemma 1 *In subgames* (q, r) *with* $q_1 < r_1$, $\underline{v}_1 < \widehat{v}_1(v_2; p, q, r) < \overline{v}_1$, *and* $\underline{v}_2 < v_2 < \overline{v}_2$, *equilibrium prices* $(\widehat{p}_A, \widehat{p}_B)$ *are determined by*

$$\widehat{p}_A - C(q) = \widehat{p}_B - C(r) = \frac{\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; q, r))] \, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1(v_2; q, r)) \, dF_2} (r_1 - q_1) \quad (16)$$

with the equilibrium curve of indifferent consumers $\widehat{v}_1(v_2; q, r) = \frac{C(r) - C(q)}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}$.

Firm A maximizes its payoffs by setting its price to find the consumer assignment $\widetilde{v}_1$ that maximizes (14). This is done by setting the price such that consumers fully internalize the cost difference between q and r. Hence, given $\widehat{p}_B$, Firm A sets its own price $\widehat{p}_A$ so that the price differential equals the cost differential $\widehat{p}_B - \widehat{p}_A = C(r) - C(q)$. The equilibrium curve of indifferent consumers is given by function $\widehat{v}_1(v_2; q, r) = \frac{C(r) - C(q)}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}$.

Regarding Firm B, the profit maximization of Firm B, given Firm A's price $\widehat{p}_A$ yields $\widehat{p}_B - C(r) = \frac{\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1)] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1) \mathrm{d}F_2} (r_1 - q_1)$. This is a version of the familiar inverse elasticity rule for Firm 2's price-cost markup (see for example Tirole 1988, p. 66). I obtain Lemma 1 by putting the two best response prices together.

Lemma 1 implies that Firm B makes positive profits. Moreover, the market shares and prices can be determined separately in the following sense. After the quality subgame, Firm A's goal is socially efficient allocation, and given Firm B's price Firm A adjusts its price to achieve that. On the other hand, because Firm B's objective is to maximize profits its best response price depends on Firm A's price in addition to the elasticity of demand.

These findings are in line with the ones from the single-attribute quality mixed duopoly model (see Lemma 1 in Laine and Ma 2017). However, two features are different. The first of these is the per-unit cost of quality: obviously in the two-attribute model the per-unit production cost of quality depends on two quality attributes instead of one. The second is that in the single-attribute quality model the price-cost markup is $\frac{1 - F(v)}{f(v)}$, which is the inverse hazard rate. In the two-attribute quality model the price-cost markup is more complicated: $\frac{\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1)] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1) \mathrm{d}F_2}$. The nominator of the price-cost markup is Firm B's demand. The denominator of the price-cost markup is the total density of the curve of indifferent consumers.

Next, I continue with the analysis by deriving how the equilibrium intercept and the equilibrium slope change with qualities (Lemma 2). These give how the curve of indifferent consumers are affected by changes in qualities, an important part when characterizing the price reaction functions for Firm A and B (Lemma 3). The price reaction functions, on the other hand, are needed for the full equilibrium characterization in the quality subgame in Subsection 3.4.2.

The complete equilibrium characterization in the quality subgame in Subsection 3.4.2 requires that I only have to consider the price reaction functions with respect to (r_1, r_2). Because Firm A's objective is to maximize social surplus, in Stage 1 Firm A's return to quality q will consist of the benefits of the consumers that buy from it. Recall that because Firm A aims at socially efficient allocation, it adjusts its price to achieve this in the price subgame (Lemma 1) and Firm A's best response price depends on the qualities determined in Stage 1.

Before these derivations I make the following simplifications for the notation. First, I re-define the equilibrium set of indifferent consumers as

$$\widehat{v}_1 (v_2; q, r) = \alpha (q, r) - \beta (q, r) v_2, \tag{17}$$

in which the intercept of the equilibrium indifference curve is defined by $\alpha (q, r) = \frac{C(r) - C(q)}{r_1 - q_1}$ and the slope of the equilibrium indifference curve is defined by $\beta (q, r) = \frac{r_2 - q_2}{r_1 - q_1}$. Both the intercept $\alpha (q, r)$ and the slope $\beta (q, r)$ of the curve of indifferent consumers $\widehat{v}_1$ are determined by the firms' quality choices.

Second, I re-write Firm B's equilibrium price as $\widehat{p}_B - C(r) = H(\alpha, \beta)(r_1 - q_1)$ in which the function $H(\alpha, \beta) : \mathbb{R}^4 \to \mathbb{R}$ is defined by $H(\alpha, \beta) \equiv \frac{\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\alpha - v_2\beta)] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta) \mathrm{d}F_2}$. Using these definitions I move on to describe how the equilibrium intercept and the equilibrium slope change with qualities.

Lemma 2 *In any subgame with qualities (q, r), the following holds in the price equilibrium:*

$$\frac{\partial \alpha (q, r)}{\partial q_i} + \frac{\partial \alpha (q, r)}{\partial r_i} = \frac{C_i(r) - C_i(q)}{r_1 - q_1} \text{ and } \frac{\partial \beta (q, r)}{\partial q_i} + \frac{\partial \beta (q, r)}{\partial r_i} = 0,$$

for $i = 1, 2$.

Lemma 2 gives the effects of the changes in the firms' qualities on the intercept of the equilibrium curve of indifferent consumers. If a firm changes either of its quality attributes, it affects the equilibrium prices of both firms in addition to the equilibrium curve of indifferent consumers. Lemma 2 says that changes in qualities changes the intercept $\alpha (q, r)$ and the slope $\beta (q, r)$ differently. It also shows that the effects of changes in the firms' qualities on the slope of the equilibrium set of indifferent consumers are equal and opposite.

Equilibrium prices $(\widehat{p}_A, \widehat{p}_B)$ relate any quality vectors to the equilibrium prices and the allocation consumers across two firms. The functional relationships between any quality vectors to the equilibrium prices and the equilibrium allocation of consumers can be written as $\widehat{p}_A (q, r)$, $\widehat{p}_B (q, r)$, and $\widehat{v}_1 (v_2; \widehat{p}, q, r)$. Total differentiation of (16) with respect to r_1 and r_2 gives the following price-reaction effects, terminology used by Barigozzi and Ma (2018):

$$\frac{\partial \widehat{p}_A}{\partial r_1} = H(\alpha, \beta) + (r_1 - q_1) \left[\frac{\partial H(\alpha, \beta)}{\partial \alpha} \frac{\partial \alpha}{\partial r_1} + \frac{\partial H(\alpha, \beta)}{\partial \beta} \frac{\partial \beta}{\partial r_1} \right] \tag{18}$$

$$\frac{\partial \widehat{p}_A}{\partial r_2} = (r_1 - q_1) \left[\frac{\partial H(\alpha, \beta)}{\partial \alpha} \frac{\partial \alpha}{\partial r_2} + \frac{\partial H(\alpha, \beta)}{\partial \beta} \frac{\partial \beta}{\partial r_2} \right] \tag{19}$$

Equations (18) and (19) show two effects. First is the effect of qualities on the intercept α and the slope β of the curve that determine the equilibrium curve of indifferent consumers. The second effect is how the equilibrium curve of indifferent consumers changes the price-cost markup $H(\alpha,\beta)$; these are terms $\frac{\partial H(\alpha,\beta)}{\partial\alpha}$ and $\frac{\partial H(\alpha,\beta)}{\beta}$.

Lemma 3 gives the full description of the strategic effects of changes in Firm B's quality on both firms' prices. I obtain these by total differentiation of (16) and by using results in the proof of Lemma 2.

Lemma 3 *From the definitions of $\widehat{p}_A$, $\widehat{p}_B$, α, β, and $H(\alpha,\beta)$, and equation (17):*

$$\begin{aligned}\frac{\partial\widehat{p}_A}{\partial r_1} = H(\alpha,\beta)\Bigg\{1+&\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}\right]\\ &[\alpha(q,r)-C_1(r)]\\ &-\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,v_2dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2f_1'(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}\right]\beta(q,r)\Bigg\}\end{aligned} \tag{20}$$

$$\begin{aligned}\frac{\partial\widehat{p}_A}{\partial r_2} = H(\alpha,\beta)\Bigg\{&\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}\right]C_2(r)\\ &+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,v_2dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha-v_2\beta)\,v_2dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}\Bigg\}\end{aligned} \tag{21}$$

$$\begin{aligned}\frac{\partial\widehat{p}_B}{\partial r_1} = C_1(r)+H(\alpha,\beta)\Bigg\{1+&\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}\right]\\ &[\alpha(q,r)-C_1(r)]\\ &-\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2f_1(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2f_1'(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}\right]\beta(q,r)\Bigg\}\end{aligned} \tag{22}$$

$$\begin{aligned}\frac{\partial\widehat{p}_B}{\partial r_2} = C_2(r)+H(\alpha,\beta)\Bigg\{&\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,dF_2}\right]C_2(r)\\ &+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha-v_2\beta)\,v_2dF_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,dF_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2f_1'(\alpha-v_2\beta)\,dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} v_2f_1(\alpha-v_2\beta)\,dF_2}\Bigg\}\end{aligned} \tag{23}$$

Lemma 3 gives the complete strategic effects of the changes in Firm B's quality on both firms' prices. It shows that these effects work in a complicated fashion. First, they depend on the production costs of the private firm's quality, the intercept, and the slope of the curve of indifferent consumers. Second, they also depend on the price-cost markup, demand, the total density of the curve of indifferent consumers and the derivative of the total density of the curve of indifferent consumers. The price reactions with respect to r_1 and r_2 of both firms are almost the same, except Firm B's equilibrium price increases at a higher rate than that of Firm A.

3.4.2 Subgame-perfect equilibrium qualities

Next, I characterize the subgame-perfect equilibrium qualities for the game with the low first quality attribute in the public firm. Given the equilibrium prices $\widehat{p}_A(q,r)$ and $\widehat{p}_B(q,r)$ in Stage 2, the continuation equilibrium payoffs for Firm A and Firm B at vector of qualities (q,r) are

$$\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,r)}\left[v_1q_1+v_2q_2-C(q)\right]\mathrm{d}F_1\right\}\mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\widehat{v}_1(v_2;q,r)}^{\overline{v}_1}\left[v_1r_1+v_2r_2-C(r)\right]\mathrm{d}F_1\right\}\mathrm{d}F_2 \tag{24}$$

and

$$\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;q,r\right)\right)\right]\left[\widehat{p}_B\left(q,r\right)-C\left(r\right)\right]\mathrm{d}F_2, \tag{25}$$

Given the subgame-perfect equilibrium prices $\widehat{p}$, equilibrium qualities $\widehat{q}$ and $\widehat{r}$ are mutual best responses:

$$\widehat{q}\equiv(\widehat{q}_1,\widehat{q}_2)=\underset{q=(q_1,q_2)}{\arg\max}\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})}\left[v_1q_1+v_2q_2-C(q)\right]\mathrm{d}F_1\right\}\mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1}\left[v_1\widehat{r}_1+v_2\widehat{r}_2-C(\widehat{r})\right]\mathrm{d}F_1\right\}\mathrm{d}F_2$$

and

$$\widehat{r}\equiv(\widehat{r}_1,\widehat{r}_2)=\underset{r=(r_1,r_2)}{\arg\max}\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right]\mathrm{d}F_2.$$

A change in Firm A's quality attribute has two effects on Firm A's payoffs. First, there is a direct effect on social surplus from those consumers who buy the good at quality q. This is $v_1q_1+v_2q_2-C(q)$. Second, the changes in either of the quality attribute changes the equilibrium prices and the market shares through the curve of indifferent consumers $\widehat{v}_1$ (Stage 2). The equilibrium prices in Stage 2 maximize social surplus. Therefore, the effect through the curve of indifferent consumers is second order because in the subgame-perfect equilibrium Lemma 1 has to hold. For Firm A, the first-order derivatives of (24) with respect to q_1 and q_2 are obtained by using Leibniz's rule and Lemma 1. Therefore, the first-order derivatives of (24) with respect to q_1 and q_2 are

$$\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})}\left[v_1-C_1(q)\right]\mathrm{d}F_1\right\}\mathrm{d}F_2 \tag{26}$$

and

$$\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})}\left[v_2-C_2(q)\right]\mathrm{d}F_1\right\}\mathrm{d}F_2. \tag{27}$$

Also, a change in Firm B's quality attribute has two effects on Firm B's profit. Qualities affect Firm B's profit (25) through three channels. First, qualities have a direct effect through the per-unit production costs and the curve of indifferent consumers (13) that affects the demand. Second, there is an effect via Firm B's own price $\widehat{p}_B$. Third, there is an effect via Firm A's price, which is captured by $\frac{\partial \widehat{p}_A}{\partial r_1}$ and $\frac{\partial \widehat{p}_A}{\partial r_2}$. Note that

$$\widehat{v}_1(q,r) \equiv \widetilde{v}_1(v_2; \widehat{p}_A(q,r), \widehat{p}_B(q,r), q, r) = \frac{\widehat{p}_B(q,r) - \widehat{p}_A(q,r)}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1} \tag{28}$$

gives the channels for the influence of r on prices.

Consider the effect via Firm B's own price $\widehat{p}_B$ first. Firm B's qualities r_1 and r_2 affect its own profit via its own equilibrium price $\widehat{p}_B$. Therefore, using Lemma 1 the effects of r_1 and r_2 on its own equilibrium price can be ignored. The first-order derivative of (25) with respect to r_1 is

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] \left[-C_1'(r)\right] \mathrm{d}F_2 \\
&+ \frac{\partial}{\partial r_1} \left\{ \int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] \mathrm{d}F_2 \right\} \times [\widehat{p}_B(\widehat{q}, r) - C(r)] \\
&+ \frac{\partial}{\partial \widehat{p}_A} \left\{ \int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r)\right] \mathrm{d}F_2 \right\} \frac{\partial \widehat{p}_A}{\partial r_1} \times [\widehat{p}_B(\widehat{q}, r) - C(r)]
\end{aligned} \tag{29}$$

and the first-order derivative of (25) with respect to r_2 is

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] [-C_2(r)] \mathrm{d}F_2 \\
&+ \frac{\partial}{\partial r_2} \left\{ \int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] \mathrm{d}F_2 \right\} \times [\widehat{p}_B(\widehat{q}, r) - C(r)] \\
&+ \frac{\partial}{\partial \widehat{p}_A} \left\{ \int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] \mathrm{d}F_2 \right\} \frac{\partial \widehat{p}_A}{\partial r_2} \times [\widehat{p}_B(\widehat{q}, r) - C(r)].
\end{aligned} \tag{30}$$

The first two rows in (30) give how private firm's quality affects its cost and demand. The last row in (30) describes the strategic effect of Firm B's quality on Firm A's price.

I obtain the first-order conditions for Firms A and B by setting (26), (27), (29), and (30) equal to zero, evaluating them at equilibrium prices, and using Lemma 1. The following proposition characterizes the equilibrium qualities.

Proposition 1 *Equilibrium qualities $(\widehat{q}, \widehat{r})$, under the assumption of $\widehat{q}_1 < \widehat{r}_1$, and the equilibrium curve of marginal consumers $\widehat{v}_1(v_2; q, r)$ solve the following equations in q_1,*

q_2, r_1, r_2:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,r)} v_1 dF_1 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;q,r\right)\right) dF_2} = C_1\left(q\right) \tag{31}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,r)} v_2 dF_1 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;q,r\right)\right) dF_2} = C_2\left(q\right) \tag{32}$$

$$\frac{\partial \widehat{p}_A}{\partial r_1} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;q,r\right)\right)\widehat{v}_1 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;q,r\right)\right) dF_2} = C_1\left(r\right) \tag{33}$$

$$\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;q,r\right)\right) v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;q,r\right)\right) dF_2} = C_2\left(r\right), \tag{34}$$

and the equilibrium curve of indifferent consumers $\widehat{v}_1\left(v_2;q,r\right)$ is given by

$$\widehat{v}_1\left(v_2;q,r\right) = \frac{C\left(r\right) - C\left(q\right)}{r_1 - q_1} - v_2\frac{r_2 - q_2}{r_1 - q_1} \tag{35}$$

in Stage 2.

I begin explaining the properties of Proposition 1 by giving the intuition in equations (31) and (32) first. Given Firm B's quality and the continuation prices, Firm A aims for the socially efficient allocation. Because Firm A's objective is to maximize social surplus, Firm A's return to quality q consists of the benefits of the consumers that buy from it. Therefore, $\widehat{q}_1$ equates the conditional average valuation of first attribute quality q_1 and the marginal per-unit cost of the first attribute quality $C_1(q)$. Similarly, $\widehat{q}_2$ equates the conditional average valuation of the second attribute quality q_2 and the marginal unit cost of the second attribute quality $C_2(q)$.

The economic intuition in equations (33) and (34) is the following. Given Firm A's qualities and the continuation equilibrium prices, Firm B's quality affects Firm A's price $\widehat{p}_A\left(q,r\right)$ in Stage 2. This is captured by $\frac{\partial \widehat{p}_A}{\partial r_1}$ and $\frac{\partial \widehat{p}_A}{\partial r_2}$. The second effect concerns the average valuation of r_1 and r_2 among the equilibrium set of indifferent consumers (denoted by integrals in (33) and (34)), the marginal unit production cost of the first quality attribute $C_1\left(r\right)$, and the marginal unit production cost of the second quality attribute $C_2\left(r\right)$. The curve of indifferent consumers is given by (35), which is obtained by using the result of equilibrium prices $\widehat{p}_A$ and $\widehat{p}_B$ in Lemma 1.

Studying the efficiency of the equilibrium outcomes requires comparison between the equilibrium outcomes and the first best. Unfortunately, this comparison turns out to be analytically intractable. I illustrate this next. Comparing equations (4)-(8) that characterize the first best to the equations that characterize the equilibrium outcome (32)-(34) shows several similarities. First, equations (4) and (5) have the similar form as equations (31) and (32). Also (8) is the same as (35). This means that the potential difference between equilibrium and the first best characterizations stems from how Firm B chooses its qualities.

This can be pinned down by subtracting the left-hand side of (33) from (6) and (34) from (7), which gives

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\alpha-v_2\beta}^{\overline{v}_1} v_1 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\alpha-v_2\beta\right)\right]\mathrm{d}F_2} - \left[\frac{\partial \widehat{p}_A}{\partial r_1} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)v_1 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}\right] \tag{36}$$

and

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\alpha-v_2\beta}^{\overline{v}_1} v_2 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\alpha-v_2\beta\right)\right]\mathrm{d}F_2} - \left[\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}\right]. \tag{37}$$

The price responses are given by Lemma 3

$$\begin{aligned}
\frac{\partial \widehat{p}_A}{\partial r_1} &= H\left(\alpha,\beta\right)\Bigg\{1+\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\alpha-v_2\beta\right)\right]\mathrm{d}F_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'\left(\alpha-v_2\beta\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}\right] \\
&\qquad \left[\alpha(q,r)-C_1\left(r\right)\right] \\
&\qquad -\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\alpha-v_2\beta\right)\right]\mathrm{d}F_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'\left(\alpha-v_2\beta\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}\right]\beta(q,r)\Bigg\} \\
\frac{\partial \widehat{p}_A}{\partial r_2} &= H\left(\alpha,\beta\right)\Bigg\{\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\alpha-v_2\beta\right)\right]\mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'\left(\alpha-v_2\beta\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}\right]C_2(r) \\
&\qquad +\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\alpha-v_2\beta\right)\right]\mathrm{d}F_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'\left(\alpha-v_2\beta\right)v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha-v_2\beta\right)\mathrm{d}F_2}\Bigg\}.
\end{aligned}$$

As shown by (36) and (37) and the price reaction functions, the comparison between general conditions becomes complicated.

However, analyzing whether the sufficient condition for efficiency in a single-attribute quality model also applies to two-attribute quality models does not necessarily require as general characterizations as in (4)-(8) and (32)-(34). I therefore rely on the two findings from the previous literature. The first finding shows that in a single-attribute quality mixed oligopoly model, the sufficient condition for efficiency is that the valuation distribution has a linear inverse hazard or inverse reverse hazard rate Laine and Ma (2017). The second finding suggests that separability in the per-unit production costs of quality plays an important role for differentiation results in a multi-attribute quality private duopoly (Barigozzi and Ma 2018).

In the following subsections I continue with the analysis by imposing additional assumptions on the valuation distributions and the per-unit cost of quality to study if these assumptions are important in a two-dimensional mixed oligopoly model too. First, I assume that the consumer quality-valuation distributions are step functions. Based on Laine and Ma (2017), in the single-attribute quality mixed oligopoly model this would result in linearity in the inverse hazard rate and therefore efficient qualities. In addition, inspired by Barigozzi and Ma (2018) I assume a simple form for the per-unit production costs of quality which allows me to study the implications of separability on differentiation and efficiency.

3.4.3 Quality differentiation and efficiency when quality-valuation density is a step function

My next results consider differentiation in the equilibrium when the quality-valuation density functions are step functions and the per-unit production cost of quality has a more specific form.

Result 2 *Suppose that f_i $i = 1, 2$ is a step function, so $f_i' = 0$ is almost everywhere and the per-unit production cost of quality is $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$.*
i) If $\theta = 0$, there is no differentiation in the second quality attribute, that is $\widehat{q}_2 = \widehat{r}_2$ in the equilibrium. ii) If $\theta \neq 0$, firms differentiate in both equality attributes, that is $\widehat{q}_2 \neq \widehat{r}_2$ in the equilibrium.

Result 2 says that if the per-unit production costs are separable and the quality-valuations follow a step function, firms differentiate only in the first quality attribute in equilibrium. But if the production costs are non-separable and the quality-valuations follow step functions, firms differentiate in both quality attributes in equilibrium.

Recall, Firm A's equations in the equilibrium characterization (equations (31) and (32)) have the same form as the respective equations in the first best characterization (equations (4) and (5)). Also, the functions that characterize the equilibrium curve of indifferent consumers and the curve of indifferent consumers in the first best are the same (equations (35) and (8) respectively). The following result considers the efficiency of the equilibrium when the per-unit production cost is separable and the quality-valuation distribution is uniform. It is based on the comparison between Firm B's first-order conditions with respect to both quality attributes and the corresponding equations in the first best characterization.

Result 3 *Suppose $q_1 < r_1$, f_i, $i = 1, 2$ is uniform, and the per-unit production cost of quality is $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$. Then, if $\theta = 0$, equilibrium qualities and market shares are the first best.*

Result 3 says that under the assumptions of uniform quality-valuations and separable per-unit production cost of quality, the equilibrium qualities and market shares are efficient. The explanation is the following. When the per-unit production of quality is separable and the quality-valuations are uniform, there is no differentiation in the first best (Result 1) and in the equilibrium (Result 2). Then under these assumptions the equations regarding the second quality attributes in the first best and the equilibrium are the same. Regarding the first quality attributes, when the second attribute qualities are equal, Firm B's first-order condition with respect r_1 given by equation (33) becomes exactly the same as it would be in one quality attribute model (Laine and Ma 2017, Proposition 1).

When the second attribute qualities are equal, H becomes the same as the inverse hazard rate of the one dimensional model, that is: $H(\widehat{v}_1) = \frac{1-F(\widehat{v}_1)}{f(\widehat{v}_1)}$, in which $\widehat{v}_1(q_1, r_1) = \frac{c(r_1)-c(q_1)}{r_1-q_1}$. Because f_1 is uniform, H is linear with $H'(\alpha) = -1$. Then I can use Proposition 2 and Remark 2 from Laine and Ma (2017), which says that the equilibrium qualities and market shares are the first best.

In the following subsection I provide further insights on these results by discussing the findings of my numerical analysis. More importantly, I use a proof by counterexample to show that linearity in the inverse hazard rate cannot be the sufficient condition for efficiency.

3.4.4 Numerical examples for $q_1 < r_1$

In this subsection, I provide numerical examples that illustrate Results 2 and 3 in addition to showing that in addition to linearity additional assumptions are needed to establish efficiency. I also use these numerical examples to compare equilibrium qualities to the numerical examples illustrating the second class of equilibria (in Section 3.5).

Consumers' valuations on two quality attributes, v_1 and v_2, are uniformly distributed on a strictly positive support $[1, 2]$ (the support has to be strictly positive so that the equilibrium qualities would be positive). The per-unit cost function of quality is quadratic: $C(q_1, q_2) = \frac{1}{2}q_1^2 + \theta q_1 q_2 + \frac{1}{2}q_2^2$.

I consider the price effects first. The first implication from assuming valuations to be uniform is the density function's property $f' = 0$. Because $f' = 0$, the second fraction in each square bracket in (20)-(23) becomes zero. The second implication of the uniform distribution assumption is that the price cost markup per unit of quality difference, $H(\alpha, \beta) = \frac{\int_{\underline{v}_2}^{\overline{v}_2} [1-F_1(\alpha - v_2\beta)] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta) \mathrm{d}F_2}$, simplifies into $H(\alpha, \beta) = 2 - \alpha + \frac{3}{2}\beta$. Using these simplifications, the strategic effects of Firm B's qualities r_1 and r_2 on equilibrium prices $\widehat{p}_A$ and $\widehat{p}_B$, from (20)-(23), become:[8]

$$\frac{\partial \widehat{p}_A}{\partial r_1} = 2 - r_1 - \theta r_2 \quad (38)$$

$$\frac{\partial \widehat{p}_A}{\partial r_2} = \frac{3}{2} - r_2 - \theta r_1 \quad (39)$$

I find the equations that characterize the equilibrium qualities in the quality subgame next. First, I substitute the price effects (38)-(39) into the equilibrium conditions (33) and (34). Then I again use the assumptions of uniform valuations and

8 $\frac{\partial \widehat{p}_B}{\partial r_1} = 2 - 2(r_1 - \theta r_2)$ and $\frac{\partial \widehat{p}_B}{\partial r_2} = \frac{3}{2} - 2(r_2 - \theta r_1)$. I have collected the details of these derivations in Appendix A.4.1.

quadratic per-unit cost of quality, and the equilibrium qualities become

$$\frac{1}{2}\left[\frac{1-\alpha^2+3\alpha\beta-\frac{7}{3}\beta^2}{1-\alpha+\frac{3}{2}\beta}\right]=q_1+\theta q_2 \tag{40}$$

$$\frac{1}{2}\left[\frac{3-3\alpha+\frac{14}{3}\beta}{1-\alpha+\frac{3}{2}\beta}\right]=q_2+\theta q_1 \tag{41}$$

$$\frac{1}{2}\left[2+\alpha-3\beta\right]=r_1+\theta r_2 \tag{42}$$

$$\frac{3}{2}=r_2+\theta r_1, \tag{43}$$

in which the curve of indifferent consumers is given by $\widehat{v}_1\left(v_1,q,r\right)=\alpha\left(q,r\right)-\beta\left(q,r\right)v_2$, with $\alpha\left(q,r\right)=\frac{\frac{1}{2}r_1^2+\theta r_1 r_2+\frac{1}{2}r_2^2-\left[\frac{1}{2}q_1^2+\theta q_1 q_2+\frac{1}{2}q_2^2\right]}{r_1-q_1}$ and $\beta\left(q,r\right)=\frac{r_2-q_2}{r_1-q_1}$.

The first two equations (40) and (41) have the same form as the respective first best equations (9) and (10). Firm A equates $\widehat{q}_1$ to the conditional average valuation of the first quality attribute q_1 with the marginal per-unit unit cost of quality (equation (40)). Similarly, $\widehat{q}_2$ equates the conditional average valuation of quality q_2 and the marginal per-unit cost of quality (equation (41)).

For Firm B, (42) equates the marginal total cost of the first quality attribute with $\frac{1}{2}\left[2+\alpha-3\beta\right]$. Equation (43) shows that when valuations are uniform, Firm B's best response is to choose the second dimension quality such that the marginal per-unit cost of the second quality attribute equals the midpoint of the whole support of v_2 that is $[1,2]$. This midpoint is 1.50.

Table 2 gives three examples of the equilibrium qualities and the social surplus for $q_1<r_1$. I have collected the results of these examples regarding differentiation in Remark 2.

Remark 2 *The numerical examples of Table 2 show that i) if $\theta=0$, there is no differentiation in second quality attribute in the equilibrium, and that ii) if $\theta\neq 0$, firms differentiate in both quality attributes in equilibrium.*
In addition to confirming the findings from Result 2, the numerical examples of Table 2 show that iii) if $\theta>0$, quality attributes are asymmetric, whereas iv) if $\theta<0$, the private firm provides superior quality and the public firm inferior quality.

In Example 4, the per-unit production cost of quality is separable. It confirms the finding i) from Result 2: under the uniform quality-valuations and when there are zero cost spillovers from the production of two quality attributes, there is no differentiation in the second quality attribute in the equilibrium. In this example, second quality attributes in the equilibrium are equal to 1.50. 1.50 is the average valuation of all v_2 consumers, and in this case the midpoint of the support of v_2. In the context of my health care provider example, this means that if the production of quality of amenities and clinical quality has no spillovers from the

TABLE 2 Equilibrium qualities and social surplus when $q_1 < r_1$.

Example		Equilibrium	qualities		Social surplus
(4) $\theta = 0$:	$\widehat{q}_1 = 1.250$	$\widehat{q}_2 = 1.500$	$\widehat{r}_1 = 1.750$	$\widehat{r}_2 = 1.500$	2.2812500
	$= q_1^{\ell *}$	$= q_2^{\ell *}$	$= q_1^{h*}$	$= q_2^{h*}$	
(5) $\theta = 0.005$:	$\widehat{q}_1 = 1.245$	$\widehat{q}_2 = 1.495$	$\widehat{r}_1 = 1.751$	$\widehat{r}_2 = 1.491$	2.2700460
	$> q_1^{\ell *}$	$< q_2^{\ell *}$	$> q_1^{h*}$	$> q_2^{h*}$	
(6) $\theta = -0.005$:	$\widehat{q}_1 = 1.255$	$\widehat{q}_2 = 1.505$	$\widehat{r}_1 = 1.749$	$\widehat{r}_2 = 1.509$	2.2925459
	$< q_1^{\ell *}$	$< q_2^{\ell *}$	$< q_1^{h*}$	$< q_2^{h*}$	

Notes. This table reports results for three different numerical examples. Valuation distribution is uniform on $[1,2]$ and the per-unit cost function of quality is quadratic. The three examples are distinguished by the value assigned to parameter θ, which is the cost spillover in producing quality attributes. Qualities refer to the equilibrium qualities for $q_1 < r_1$. Social surplus refers to the maximum social surplus that the equilibrium qualities yield.

production of qualities, in the equilibrium the firms provide goods that are differentiated in the amenity dimension and provide equal clinical qualities.

Examples 5 and 6 give numerical examples with the uniform quality-valuations and with the non-zero cost spillovers from the production of the two quality attributes. It confirms the finding ii) from Result 2: when there are non-zero cost spillovers from the quality production, there is differentiation in both dimensions. In the context of the health care provider example, if the per-unit production cost of quality has spillovers from the production of amenities and clinical quality, the firms provide products that are differentiated in both amenities and clinical qualities.

The new finding from these examples relative to Result 2 considers how quality attributes are differentiated relative to each other when the there is differentiation in both dimensions. In Example 5, the cost spillover from the production of two quality attributes is positive. In this equilibrium, differentiation is asymmetric: Firm B produces a higher first attribute quality but a lower second attribute quality than Firm A. In Example 6 the cost spillover from the production of two quality attributes is negative. In this example, Firm B's product is of a superior quality, and Firm A's is inferior.

The following Remark 3 collects these numerical findings regarding efficiency in addition to the results regarding differentiation relative to the other firm's quality attribute.

Remark 3 *The numerical examples of Table 2 show that i) if $\theta = 0$, the equilibrium coincides with the first best, and that ii) if $\theta \neq 0$, the equilibrium's quality attributes in each dimension are higher or lower than the first best. Also iii) if $\theta > 0$, the asymmetric equilibrium quality attributes can be lower or higher than in the first best, whereas iv) if $\theta < 0$, both the inferior and superior equilibrium quality attributes are below the first-best.*

By comparing the equilibrium quality values in Example 4 to the ones in Example

1 in Table 1, under assumptions of uniform valuations and these particular costs the equilibrium qualities coincide with the first best thus confirming the finding in Result 3.

However, the numerical examples above show by counterexample that linearity in the inverse hazard rate cannot be the sufficient condition for efficiency. If that were the case, all equilibrium qualities in Examples 4-6 would be efficient. Examples 5 and 6 give numerical examples with the uniform quality-valuations and with the non-zero cost spillovers from the production of the two quality attributes. Recall that uniform quality-valuation is the simplest valuation distribution with a linear inverse hazard rate. The qualities are inefficient in Examples 5 and 6 and thus linearity in the inverse hazard rate cannot be the sufficient condition for efficiency.

In Examples 5 and 6, the equilibrium qualities and the market shares are inefficient, which confirms Remark 2. More important, these examples illustrate how quality attributes can be differentiated relative to each other. The example shows that the first quality attribute of Firm A is higher than the first best, and the second quality attribute is lower than the first best. Both Firm B's quality attributes are higher than the first best. In Example 6, all quality attributes in both firms are lower than the first best. This example also yields the highest social surplus.

3.5 Equilibria with high first quality attribute at the public firm

In this section, I characterize the second class of equilibria with $q_1 > r_1$. In my health care provider example this would represent a case in which Firm A offers higher quality amenities than Firm B. As I discussed in Subsection 3.2.3, analyzing these two classes separately is important because the two firms have different objectives; the welfare payoffs in these two classes are not isomorphic. I illustrate this and discuss its implications in Subsection 3.5.4.

I begin by solving the subgame-perfect equilibrium prices and then continue by solving the subgame-perfect equilibrium qualities. As many of the steps for the derivations and proofs follow the ones in Section 3.4, I have omitted them and provide only those steps that are relevant for my analysis.

3.5.1 Subgame-perfect equilibrium prices

Consider a subgame (q, r) for $q_1 > r_1$ in Stage 2, in which q and r are vectors of qualities. A function

$$\widetilde{v}_1 \left(v_2; p_A, p_B, q, r\right) = \frac{p_A - p_B}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}, \tag{44}$$

defines all consumers who are indifferent between buying from Firm A and Firm B. For the equilibrium characterization of $q_1 > r_1$, I re-write the curve of indifferent consumers such as in (44). In this case a consumer (v_1', v_2) buys from Firm A if and only if $v_1' > \widetilde{v}_1(v_2; p, q, r)$. In Figure 1, this would be the curve of consumers with valuations above $\widetilde{v}_1$. A consumer (v_1', v_2) buys from Firm B if and only if $v_1' < \widetilde{v}_1(v_2; p, q, r)$ which is the curve of consumers with valuations below $\widetilde{v}_1$ in Figure 1. I re-write the curve as (44) to clarify the analysis, and it does not affect my results.

Firm A's payoffs are

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\widetilde{v}_1(v_2;p,q,r)} [v_1 r_1 + v_2 r_2 - C(r)] \, \mathrm{d}F_1 \right\} \mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\widetilde{v}_1(v_2;p,q,r)}^{\overline{v}_1} [v_1 q_1 + v_2 q_2 - C(q)] \, \mathrm{d}F_1 \right\} \mathrm{d}F_2, \tag{45}$$

in which (p_A, p_B) is the vector of prices and $\widetilde{v}_1$ is defined by (44).

Firm B's payoffs are

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\widetilde{v}_1(v_2;p,q,r)} [p_B - C(r)] \, \mathrm{d}F_1 \right\} \mathrm{d}F_2 = \int_{\underline{v}_2}^{\overline{v}_2} F_1(\widetilde{v}_1(v_2; p, q, r)) [p_B - C(r)] \, \mathrm{d}F_2, \tag{46}$$

in which (p_A, p_B) is the vector of prices, and $\widetilde{v}_1$ is defined by (44).

Equilibrium prices $(\widehat{p}_A, \widehat{p}_B)$ are best responses against each other. These are characterized by the following:

Lemma 4 *In subgames (q, r) with $q_1 > r_1$, $\underline{v}_1 < \widehat{v}_1(v_2; p, q, r) < \overline{v}_1$, and $\underline{v}_2 < v_2 < \overline{v}_2$, equilibrium prices $(\widehat{p}_A, \widehat{p}_B)$ are*

$$\widehat{p}_A - C(q) = \widehat{p}_B - C(r) = \frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2; \widehat{p}, q, r)) \, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1(v_2; \widehat{p}, q, r)) \, dF_2} (q_1 - r_1) \tag{47}$$

with the equilibrium curve of indifferent consumers $\widehat{v}_1(v_2; q, r) = \frac{C(q) - C(r)}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}$.

The properties of Lemma 4 are parallel to those in Lemma 1. Firm A sets its price to find the consumer assignment that maximizes (45) setting the price difference equal to the cost difference: $\widehat{p}_A - \widehat{p}_B = C(q) - C(r)$. The profit maximization of Firm B yields $\widehat{p}_B - C(r) = \frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2;\widehat{p},q,r)) \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1(v_2;\widehat{p},q,r)) \mathrm{d}F_2} (q_1 - r_1)$.

Next, I continue with the analysis by deriving the price reaction functions for Firm A and B. As above, I use these price reaction functions for the equilibrium characterization in the quality subgame. Again, because Firm A aims at socially efficient allocation, I only have to consider the price reaction functions with respect to r.

Before these derivations, I simplify the notation by re-defining the equilibrium curve of indifferent consumers as $\widehat{v}_1(v_2; q, r) = \gamma(q, r) - \delta(q, r) v_2$, in which the intercept of the equilibrium indifference curve is $\gamma(q, r) = \frac{C(q) - C(r)}{q_1 - r_1}$, and the slope of the equilibrium indifference curve is $\delta(q,r) = \frac{q_2 - r_2}{q_1 - r_1}$. I rewrite Firm B's equilibrium price as

$$\widehat{p}_B - C(r) = G(\gamma,\delta)(q_1 - r_1). \tag{48}$$

in which the function $G(\gamma,\delta) : \mathbb{R}^4 \to \mathbb{R}$ is defined by

$$G(\gamma,\delta) \equiv \frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}. \tag{49}$$

The nominator of the price-cost markup G is Firm B's demand. The denominator of the price-cost markup is the total density of the curve of indifferent consumers. With the simplifying notation, the equilibrium prices in Lemma 4 become

$$\widehat{p}_A - C(q) = \widehat{p}_B - C(r).$$

The equilibrium prices $(\widehat{p}_A, \widehat{p}_B)$ relate any quality vectors to the equilibrium prices and the allocation of consumers across two firms. These functional relationships can be written as $\widehat{p}_A(q,r)$, $\widehat{p}_B(q,r)$, and $\widehat{v}_1(v_2; \widehat{p}, q, r)$. The following summarizes how the prices react to changes in the private firm's qualities (r_1, r_2):

Lemma 5 *From the definitions of* $\widehat{p}_A$ *and* $\widehat{p}_B$ *the price reaction functions are*

$$\begin{aligned}\frac{\partial \widehat{p}_A}{\partial r_1} = G(\gamma,\delta)\Bigg\{ -1 + &\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\right] \\ &[\gamma(q,r) - C_1(r)] \\ + &\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\right]\delta(q,r)\Bigg\}\end{aligned} \tag{50}$$

$$\begin{aligned}\frac{\partial \widehat{p}_A}{\partial r_2} = G(\gamma,\delta)\Bigg\{ - &\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\right] C_2(r) \\ + &\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\, v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\Bigg\}\end{aligned} \tag{51}$$

$$\begin{aligned}\frac{\partial \widehat{p}_B}{\partial r_1} = C_1(r) + G(\gamma,\delta)\Bigg\{ -1 + &\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\right] \\ &[\gamma(q,r) - C_1(r)] \\ + &\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\right]\delta(q,r)\Bigg\}\end{aligned} \tag{52}$$

$$\begin{aligned}\frac{\partial \widehat{p}_B}{\partial r_2} = C_2(r) + G(\gamma,\delta)\Bigg\{ - &\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\, dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\right] C_2(r) \\ + &\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\, dF_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\, v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\, dF_2}\Bigg\}.\end{aligned} \tag{53}$$

Lemma 5 gives the strategic effects of changes in Firm B's quality on both firms' prices. The strategic effects work in a complicated fashion. Such as in the first equilibrium class, they depend on the production costs of Firm B's quality, in addition to the intercept and the slope of the equilibrium curve of indifferent consumers. The strategic effects also depend on Firm B's price-cost markup, demand, the total density of the curve of indifferent consumers, and the derivative of the total density of the set of indifferent consumers. I use these price reaction functions for the complete characterization of the subgame-perfect qualities in the following subsection.

3.5.2 Subgame-perfect equilibrium qualities

Next, I characterize the subgame-perfect equilibrium qualities for the game with a high first quality attribute in the public firm.

Given the subgame-perfect equilibrium prices $\widehat{p}_A$ and $\widehat{p}_B$ in Stage 2, the equilibrium qualities $\widehat{q}$ and $\widehat{r}$ are mutual best responses:

$$\widehat{q} \equiv (\widehat{q}_1, \widehat{q}_2) = \underset{q=(q_1,q_2)}{\arg\max} \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})} [v_1 \widehat{r}_1 + v_2 \widehat{r}_2 - C(\widehat{r})] \, \mathrm{d}F_1 \right\} \mathrm{d}F_2$$
$$+ \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1} [v_1 q_1 + v_2 q_2 - C(q)] \, \mathrm{d}F_1 \right\} \mathrm{d}F_2$$

and

$$\widehat{r} \equiv (\widehat{r}_1, \widehat{r}_2) = \underset{r=(r_1,r_2)}{\arg\max} \int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2; \widehat{q}, r)) [\widehat{p}_B(\widehat{q}, r) - C(r)] \, \mathrm{d}F_2.$$

I apply the same method as in Section 3.4.2 to characterize the equilibrium qualities. Firm A chooses its quality attributes to maximize the surplus of those consumers above $\widehat{v}_1$. Changes in q_1 and q_2 in Firm A's payoff affect only the second integral because the effect via the first integral is secondary because of the Envelope Theorem. Moreover, changes in r have two effects. The first effect is the effect on the surplus of the marginal consumer $v_1 r_1 + v_2 r_2 - C(r)$. Second is the effect on Firm A's price $\widehat{p}_A(q, r)$ in Stage 2. Any effect on the marginal consumer is second order because of the Envelope Theorem. For Firm B, I obtain the first order conditions by taking the first-order derivatives with respect to both quality attributes r_1 and r_2 and using Lemma 4. The first-order conditions give the following result on the equilibrium qualities and allocation.

Proposition 2 *Equilibrium qualities $(\widehat{q}, \widehat{r})$, under the assumption of $\widehat{r}_1 < \widehat{q}_1$, and the equilibrium curve of marginal consumers $\widehat{v}_1(v_2; q, r)$ must satisfy the following equations in q_1, q_2, r_1, r_2:*

$$\frac{\partial \widehat{p}_A}{\partial r_1} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1(v_2; q, r)) \widehat{v}_1 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1(v_2; q, r)) dF_2} = C_1(r) \tag{54}$$

$$\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1(v_2; q, r)) v_2 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\widehat{v}_1(v_2; q, r)) dF_2} = C_2(r) \tag{55}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \int_{\widehat{v}_1(v_2;q,r)}^{\overline{v}_1} v_1 dF_1 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; q, r))] dF_2} = C_1(q) \tag{56}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \int_{\widehat{v}_1(v_2;q,r)}^{\overline{v}_1} v_2 dF_1 dF_2}{\int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; q, r))] dF_2} = C_2(q), \tag{57}$$

and the equilibrium curve of indifferent consumers $\widehat{v}_1(v_2; q, r)$ is given by

$$\widehat{v}_1(v_2; q, r) = \frac{C(q) - C(r)}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}. \tag{58}$$

in Stage 2.

The properties of the equations in Proposition 2 are similar to those in Proposition 1. Given Firm A's quality and the continuation equilibrium prices, Firm B's quality affects Firm A's price $\widehat{p}_A(q,r)$ in Stage 2, which is captured by $\frac{\partial \widehat{p}_A}{\partial r_1}$ and $\frac{\partial \widehat{p}_A}{\partial r_2}$. The second effect concerns the average valuation of r_1 and r_2 among the equilibrium curve of indifferent consumers (denoted by integrals in (54) and (55)), and the per-unit-costs $C_1(r)$ and $C_2(r)$. These are the first two equations, (54) and (55).

Given Firm B's quality and the continuation equilibrium prices, Firm A's return from qualities q_1 and q_2 consists of its own consumers. Then, $\widehat{q}_1$ equates the conditional average valuation of quality q_1 and the marginal cost $C_1(q)$. Similarly, $\widehat{q}_2$ equates the conditional average valuation of quality q_2 and the marginal cost $C_2(q)$. These are the third and fourth equations, (56) and (57). I obtain the curve of indifferent consumers $\widehat{v}_1(v_2; q, r)$ by using the result of equilibrium prices $\widehat{p}_A$ and $\widehat{p}_B$ in Lemma 5.

Despite the similarities between the general characterizations of the subgame equilibrium qualities in Propositions 1 and 2, characterizing both class of equilibria ($q_1 < r_1$ and $q_1 > r_1$) is important. This is because the equilibria may yield different equilibrium outcomes and social surplus depending on whether the public firm offered the lower or higher first quality attribute than the private firm.

These differences appear already in the single-attribute model (see Sections 4.3 and 5.1 in Laine and Ma 2017). Note that some quality-valuation distributions may have linear inverse hazard (or linear inverse reverse hazard), but not necessarily both. Thus, the results in Laine and Ma (2017) imply that by entering the correct quality segment the public firm can restore efficiency. Uniform distribution has both linear inverse hazard and linear inverse reverse hazard. For example, for the triangular distribution $f(v) = 2v$, $G(v)$ is linear, but $H(v)$ is not. For the reverse triangular distribution $f(v) = 2 - 2v$, $H(v)$ is linear, but $G(v)$ is not. Laine and Ma (2017) show that generally the two equilibrim classes yield different social surpluses and therefore highlight the importance of policy design: it is important that the public firm enters the correct market segment (high or low quality attribute).

As shown in the examples in Subsection 3.4.4, unlike in Laine and Ma (2017) the linearity in the inverse hazard rate is not a sufficient condition for efficiency in a two-attribute quality mixed duopoly. In the following subsections I complete the analysis by showing that the same applies also for the second equilibrium class. I also give numerical examples that show that if the equilibrium qualities are inefficient, the equilibrium qualities in addition to the social surplus are different depending on whether the public firm offers a lower or higher first quality attribute than the private firm. This suggests implications to the competition policy in a similar fashion as in the single-attribute model.

3.5.3 Quality differentiation and efficiency when quality-valuation density is a step function

As in the previous sections, I impose additional assumptions on the quality-valuation distributions and the unit cost of quality, provide the results on differentiation and efficiency, and then move on to my numerical examples.

My first result considers differentiation in the equilibrium by assuming that the quality-valuation density functions are step functions and the per-unit production cost of quality is separable. Proof of Result 4 follows closely the proof of Remark 2 and is omitted.

Result 4 *Suppose $q_1 > r_1$, f_i, $i = 1, 2$ is a step function, so $f_i' = 0$ is almost everywhere and the per-unit production cost of quality is $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$.*
i) If $\theta = 0$, there is no differentiation in the second quality attribute, that is $\widehat{q}_2 = \widehat{r}_2$ in the equilibrium. ii) If $\theta \neq 0$, firms differentiate in both equality attributes, that is $\widehat{q}_2 \neq \widehat{r}_2$ in the equilibrium.

Similarly to Result 2 for the first equilibrium class, if the per-unit production cost of quality is separable and the quality-valuations follow step functions, firms differentiate only in the first quality attribute in the second equilibrium class (finding i) in Result 4). Also the finding ii) in Result 4 is the same, if the per-unit production cost of quality is non-separable and the quality-valuations follow step functions, there will be differentiation in two quality attributes in the second equilibrium class as well.

The following result considers the efficiency of the equilibrium when the per-unit production cost is separable and the quality-valuation distribution is uniform. Because Firm A's equations for the equilibrium characterization (equations (56) and (57)) are the same as the respective equations in the first best characterization (equations (6) and (7)) and the functions that characterize the equilibrium curve of indifferent consumers and the curve of indifferent consumers in the first best are the same (equations (4) and (5) respectively), I can obtain the result by comparing Firm B's first-order conditions (54) and (55) with respect to both quality attributes to the corresponding equations in the first best characterization (equations (58) and (8).

Result 5 *Suppose $q_1 > r_1$, f_i, $i = 1, 2$ is uniform and the per-unit production cost of quality is $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$. If $\theta = 0$, the equilibrium qualities are the first best.*

When the quality-valuations are uniform and the per-unit production cost of quality is separable, the result in Laine and Ma (2017) also applies to the second equilibrium class: the equilibrium qualities and market shares coincide with the first best. As above, under the simplifying assumptions on the quality-valuations and the per-unit costs the second quality attributes in the first best and the equilibrium are the same. The equations regarding the first quality attributes become the same as they would be in the single-quality attribute model (see Proposition 5 in Laine and Ma 2017). Thus because the second quality attributes are the same, the equations regarding the second quality attributes in the first best and the equilibrium will become the same.

Under these assumptions for the second equilibrium class, G becomes the inverse reverse hazard rate $G(\widehat{v}_1) = \frac{F(\widehat{v}_1)}{f(\widehat{v}_1)}$. Because f_i is uniform, G is linear with $G'(\widehat{v}_1) = 1$. Then I can use Proposition 5 and Remark 4 from Laine and Ma (2017), which says that the equilibrium qualities and market shares are the first best.

In the following subsection I provide further insights on these results by discussing the findings of my numerical analysis. I use proof by counterexample to show that linearity in the inverse hazard rate cannot be the sufficient condition for efficiency. Last, I illustrate how the two equilibrium classes may yield different equilibrium outcomes and social surpluses.

3.5.4 Numerical examples for $q_1 > r_1$

This subsection provides numerical examples that illustrate Results 4 and 5. More importantly, I also compare the equilibrium outcomes of these examples to the ones in Tables 1 and 2. These results highlight the importance of two classes of equilibria and illustrate the variety of quality segmentations in mixed oligopolies.

Consumers' valuations on two quality attributes v_1 and v_2 are uniformly distributed on a strictly positive support $[1, 2]$. The per-unit production cost of quality is quadratic: $C(q_1, q_2) = \frac{1}{2}q_1^2 + \theta q_1 q_2 + \frac{1}{2}q_2^2$.

I consider the price-reaction functions first. As a result of assuming uniform valuations, the price-cost markup per unit of quality difference becomes $G(\gamma, \delta) = \gamma - \frac{3}{2}\delta - 1$, and the strategic effects of Firm B's qualities r_1 and r_2 on equilibrium price $\widehat{p}_A$ simplify into:[9]

$$\frac{\partial \widehat{p}_A}{\partial r_1} = 1 - r_1 - \theta r_2 \tag{59}$$

$$\frac{\partial \widehat{p}_A}{\partial r_2} = \frac{3}{2} - r_2 - \theta r_1. \tag{60}$$

9 $\frac{\partial \widehat{p}_B}{\partial r_1} = 1 - 2(r_1 - \theta r_2)$ and $\frac{\partial \widehat{p}_B}{\partial r_2} = \frac{3}{2} - 2(r_2 - \theta r_1)$. For details of the derivations, see Appendix A.5.1.

I substitute the price-reaction effects (59) and (60) for the (54) and (55) and use the assumption of uniform valuations, and the equilibrium in the quality subgame with $q_1 > r_1$ becomes:

$$\frac{1}{2}\left[1+\gamma-\frac{3}{2}\delta\right] = r_1 + \theta r_2 \tag{61}$$

$$\frac{3}{2} = r_2 + \theta r_1 \tag{62}$$

$$\frac{1}{2}\left[\frac{2-\gamma^2+3\gamma\delta-\frac{7}{3}\delta^2}{2-\gamma+\frac{3}{2}\delta}\right] = q_1 + \theta q_2 \tag{63}$$

$$\frac{1}{2}\left[\frac{3-3\gamma+\frac{14}{3}\delta}{2-\gamma+\dfrac{3}{2}\delta}\right] = q_2 + \theta q_1, \tag{64}$$

in which the curve of the indifferent consumers is given by $\widehat{v}_1\left(v_1, q, r\right) = \gamma\left(q, r\right) - \delta\left(q, r\right) v_2$, with $\gamma\left(q, r\right) = \frac{\frac{1}{2}q_1^2+\theta q_1 q_2+\frac{1}{2}q_2^2-\left[\frac{1}{2}r_1^2+\theta r_1 r_2+\frac{1}{2}r_2^2\right]}{q_1-r_1}$, and $\delta\left(q, r\right) = \frac{q_2-r_2}{q_1-r_1}$.

Table 3 gives three examples of the equilibrium qualities and social surplus for $q_1 > r_1$. The first two columns in Table 3 present the results on r first because these examples consider the equilibria with high first quality attribute in the public firm. The following remark collects the conclusions of these examples regarding differentiation in addition to how products are differentiated relative to the other firm's quality attributes.

TABLE 3 Equilibrium qualities and social surplus when $q_1 > r_1$.

Example		Equilibrium	qualities		Social surplus
(7) $\theta = 0$:	$\widehat{r}_1 = 1.250$	$\widehat{r}_2 = 1.500$	$\widehat{q}_1 = 1.750$	$\widehat{q}_2 = 1.500$	2.2812500
	$= q_1^{\ell*}$	$= q_2^{\ell*}$	$= q_1^{h*}$	$= q_2^{h*}$	
(8) $\theta = 0.005$:	$\widehat{r}_1 = 1.251$	$\widehat{r}_2 = 1.494$	$\widehat{q}_1 = 1.745$	$\widehat{q}_2 = 1.490$	2.2700453
	$> q_1^{\ell*}$	$< q_2^{\ell*}$	$> q_1^{h*}$	$> q_2^{h*}$	
(9) $\theta = -0.005$:	$\widehat{r}_1 = 1.249$	$\widehat{r}_2 = 1.506$	$\widehat{q}_1 = 1.755$	$\widehat{q}_2 = 1.510$	2.2925466
	$< q_1^{\ell*}$	$> q_2^{\ell*}$	$< q_1^{h*}$	$< q_2^{h*}$	

Notes. This table reports results for three different numerical examples. Valuation distribution is uniform on $[1, 2]$, and the per-unit cost function of quality is quadratic. The three examples are distinguished by the value assigned to parameter θ, which is the cost spillover in producing quality attributes. Qualities refer to the equilibrium qualities for $q_1 < r_1$. Social surplus refers to the maximum social surplus that the equilibrium qualities yield.

Remark 4 *The numerical examples of Table 3 show that i) if $\theta = 0$, there is no differentiation in the second quality attribute in the equilibrium, and that ii) if $\theta \neq 0$, firms differentiate in both quality attributes in equilibrium.*
In addition to confirming the findings from Result 3, the numerical examples of Table 3 show that iii) if $\theta > 0$, quality attributes are asymmetric, whereas iv) if $\theta < 0$, the public firm provides superior quality and the private firm inferior quality, and v) the direction of differentiation in iii) and iv) is the same as in the first equilibrium class.

In Example 7, the per-unit production cost of quality is separable, and in Examples 8 and 9 the per-unit production cost of quality is non-separable. These examples confirm the findings i) and ii) from Result 4: under the uniform quality-valuations and when there are zero cost spillovers from the production of two quality attributes, there is no differentiation in the second quality attribute in equilibrium. If the per-unit production cost of quality is non-separable, there is differentiation in both dimensions in both examples.

Again, the new finding from these examples relative to Result 3 considers how quality attributes are differentiated relative to each other when the there is differentiation in both dimensions. In Example 8 the cost spillover from the production of the two quality attributes is positive. In this example the differentiation in equilibrium is asymmetric. In Example 9 the cost spillover from the production of the two quality attributes is negative. Now Firm A produces a product of superior quality, and Firm B's product is inferior.

The numerical examples allow me to compare the direction of differentiation in the first equilibrium class to the direction of the differentiation in the second equilibrium class. In the examples, the differentiation goes in the same direction in two classes: if $\theta > 0$, similarly to the first equilbrium class, the quality attributes are asymmetric in the second equilibrium class. Similarly, if $\theta < 0$, similarly to the first equilibrium class, the *private* firm provides superior quality and the *public* firm inferior quality $q_1 < r_1$. The numerical examples allow me to compare the direction of differentiation in the first equilibrium class to the direction of the differentiation in the second equilibrium class. In the examples, the differentiation goes in the same direction in two classes: if $\theta > 0$, similarly to the first equilbrium class, the quality attributes are asymmetric in the second equilibrium class. Similarly, if $\theta < 0$, similarly to the first equilibrium class, the *private* firm provides superior quality and the *public* firm inferior quality $q_1 < r_1$.

The following remark collects my findings from my numerical examples regarding the efficiency of the equilibrium.

Remark 5 *The numerical examples of Table 3 show that i) if $\theta = 0$, the equilibrium coincides with the first best, and that ii) if $\theta \neq 0$, the equilibrium is not the first best. Also, iii) if $\theta > 0$, the asymmetric equilibrium quality attributes can be lower or higher than in the first-best, whereas iv) if $\theta < 0$, both the inferior and superior quality attributes can be lower or higher than in the first-best.*

Comparing the equilibrium outcomes of Example 7 to the first best shows that when the quality-valuations are uniform and the per-unit production cost of quality is separable, the equilibrium qualities are efficient. This confirms Result 5. Moreover, when the spillovers from quality production costs are present, the qualities are inefficient, and firms have an incentive to differentiate in all quality dimensions. When quality spillovers from production is positive, the first quality attribute of Firm A is higher than the first best, and the second quality is lower than the first best. Both of Firm B's quality attributes are higher than the first best.

Recall that Examples 5, 6, 8, and 9 are all counterexamples of the quality-valuation distributions that give linear inverse hazard and inverse reverse hazard in a single-attribute model but give inefficient equilibrium outcomes when there are two quality attributes. My examples show that similar properties for the quality-valuation as those in the single-attribute quality model cannot be the sufficient condition for efficiency if qualities consist of more than one attribute. Results 3 and 5 suggest this additional assumption to be separability in the per-unit production cost of quality. This is also confirmed by Examples 1, 4, and 7.

Last, but not least, the following remark summarizes results regarding the differences in equilibrium qualities and social surplus from the two classes of equilibria.

Remark 6 *The numerical examples of Tables 2 and 3 show i) if $\theta = 0$, the equilibrium qualities are the mirror images of each other and the social surplus is the same in the two classes of equilibria, and that ii) $\theta \neq 0$, the equilibrium qualities and the social surplus are all different in the two classes of equilibria.*

All examples illustrate the inherent feature in mixed oligopolies: because the two firms have different objectives, the two equilibrium classes may yield different outcomes. Recall that the examples of the two equilibrium classes have otherwise the same parameter values but with a different order of first attribute qualities. The examples show that when the equilibrium qualities are not efficient, the resulting equilibrium qualities and social surplus in a mixed duopoly can be different depending on the production technology that is used, or if the public firm enters the lower or higher segment of the first quality dimension. In Examples 5, 6, 8, and 9 social surplus is the highest when there are negative spillovers of quality production on costs. Interestingly the examples also show that when there are positive spillovers in quality production the social surplus is higher for the equilibrium class with low first quality attribute in the public firm ($q_1 < r_1$), but when there are negative spillovers of quality production on costs the social surplus is higher for the equilibrium class with high first quality attribute in the public firm ($q_1 > r_1$).

Last, note that differences arise even though the shape of the quality-valuation distribution is very symmetric. Results from Laine and Ma (2017), however, suggest that other distributions for the quality-valuation with different shapes may affect the equilibrium outcome and even amplify the differences between two equilibrium classes. On overall, whether the numerical findings presented above are more generalizable is a question that I will continue to explore in the subsequent versions of this paper.

3.6 Conclusions

I have studied a mixed duopoly model in which a public and a private firm compete on prices and on two-attribute qualities. Consumers have different quality-valuations that follow independent valuation distributions. The public firm maximizes social surplus, and the private firm maximizes profit. The per-unit production cost of quality is an increasing and a convex function.

I have characterized the first best and the equilibrium in a two-dimensional mixed duopoly model. After characterizing the first best and the equilibrium, I made additional assumptions on the consumer valuation distributions and the per-unit production costs. I showed that if the per-unit cost of producing two quality attributes is separable in the two quality attributes, there is only differentiation in one quality dimension. If the per-unit cost of quality production is not separable, there will be differentiation in both quality attributes. By adding more structure to the model I studied how the assumptions on the quality-valuations and the per-unit production cost of quality affect the results regarding multi-dimensional differentiation and the efficiency of the equilibrium outcome.

I found that unlike in the single-attribute quality mixed duopoly model in Laine and Ma (2017), the linearity in the inverse hazard or inverse reverse hazard is not a sufficient condition for efficiency when the quality consists of two attributes. My results showed that even with the simplest valuation distribution satisfying the linearity assumption, that is the uniform distribution, the equilibrium outcome in a two-attribute quality mixed duopoly is not always efficient. In contrast, I showed that when the quality-valuations are uniform (and independent), the market outcome is efficient if the per-unit production cost of the two quality attributes is separable.

Regarding competition policy and how my results relate to the equilibria in private duopolies, recall, that if the market consists of two private firms the equilibrium qualities are always inefficient. My results suggest that in some cases equilibria in a two-attribute quality mixed duopoly might even be efficient. These findings give support to the standard suggestion in the mixed oligopoly literature that the social surplus would be improved if one of the private firms were taken over by a social-surplus maximizing public firm (Cremer et al. 1991; Grilo 1994; Laine and Ma 2017).

The analysis in this paper can be extended in several ways. First, I used numerical examples to compare the equilibrium outcomes to the first-best qualities under assumptions of uniform valuation distribution, as well as separable and non-separable unit production cost of qualities. This proof by counterexample showed that unlike in a one quality attribute mixed duopoly, linearity in the inverse hazard or inverse reverse hazard cannot be the sufficient condition for efficiency.

I assume that quality consists of two-attributes, so extending the model to consider N attributes should be fairly straightforward. Now the function that determines the curve of indifferent consumers in (13) should consist of the sum of N valuations. Once the curve of indifferent consumers is specified, the usual optimization steps can be taken to characterize the first best and the equilibrium. Also extending the model to allow correlated valuations (complements or substitutes) would be interesting too.

The characterizations above have assumed the existence of the subgame-perfect equilibria, and thus the analysis focused on the necessary first-order conditions for the optimality. More detailed research on the existence of the equilibrium in a vertical two-attribute quality mixed oligopoly model by using the general model would be a natural next step.

Lastly, the vast majority of the analysis has focused on valuation functions that are either step functions or uniform, as well as on quadratic production costs. More general results using the general model would be an obvious important next step. Moreover, the general model could be used in studying how the efficiency is affected by different forms for the valuation distributions and the per-unit production cost functions.

3.A Appendix: Proofs and derivations

This appendix contains the proofs of the results reported in the main text. Appendix 3.A.1 contains the proofs of the results in Section 3.3 Appendix 3.A.2 collects the proofs of the results in Section 3.4 Appendix 3.A.3 collects the proofs of the results in Section 3.5 Appendix 3.A.4 contains the derivations of the equations for my numerical examples. The conditions for existence are discussed in Appendix 3.A.5. Appendix 3.A.6 discusses the protocol of my numerical simulations in detail.

I use the following notation in the proofs: v denotes the vector of valuations and v_i is each valuation that follows the distribution function F_i with the corresponding density f_i, $i = 1, 2$. I sometimes use $\mathrm{d}F_i$ to denote the corresponding density of F_i. Then, the derivative of the density is denoted either by f_i' or $\mathrm{dd}F_i$.

I also use the following simplifying notation for the partial derivatives of the per-unit production cost of quality with respect to each quality in each dimension: $C_1(q) \equiv \frac{\partial C(q)}{\partial q_1}$, $C_2(q) \equiv \frac{\partial C(q)}{\partial q_2}$. Also, in some of my results I use the following form for the per-unit cost function $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$, with $c_1(q) \equiv \frac{\partial c(q_1)}{\partial q_1}$, $c_2(q) \equiv \frac{\partial c(q_2)}{\partial q_2}$.

3.A.1 Proofs of the results in Section 3.3

Derivation of equations (4)-(8)

Here, I provide the steps for the derivations of equations (4)-(7). Equations (4)-(7) together with the function given by (8) characterize the first best qualities and an assignment of consumers, denoted by $(q^{\ell*}, q^{h*}, v_1^*)$. Last, I discuss the conditions that should be satisfied for the first-best in (3) to have a unique maximum.

As described in the main text, an allocation consists of two vectors of product qualities $\left(q^\ell, q^h\right)$, one vector each at Firm ℓ and Firm h, and an assignment of consumers across the two firms that will be given by $\check{v}_1$. The social surplus of an allocation is the sum of the consumer surplus and firms' profit:

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\check{v}_1} \left[v_1 q_1^\ell + v_2 q_2^\ell - C(q^\ell) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\check{v}_1}^{\overline{v}_1} \left[v_1 q_1^h + v_2 q_2^h - C(q^h) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2, \tag{65}$$

in which $q_1^\ell < q_1^h$. Social planner chooses quality vectors q^ℓ and q^h in addition $\check{v}_1$ to maximize social surplus (65).

I begin by finding the first-order conditions with respect to qualities. I differentiate (65) with respect to q_1^ℓ and set the first-order derivative to zero to obtain

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\check{v}_1} \left[v_1 - C_1\left(q^\ell\right) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 = 0. \tag{66}$$

Second, I differentiate (65) with respect to q_2^ℓ and set the first-order derivative to zero to obtain

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\check{v}_1} \left[v_2 - C_2\left(q^\ell\right) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 = 0. \tag{67}$$

Third, I differentiate (65) with respect to q_1^h and set the first-order derivative to zero to obtain

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\check{v}_1}^{\overline{v}_1} \left[v_1 - C_1\left(q^h\right) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 = 0. \tag{68}$$

Fourth, I differentiate (65) with respect to q_2^h and set the first-order derivative to zero to obtain

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\check{v}_1}^{\overline{v}_1} \left[v_2 - C_2\left(q^h\right) \right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 = 0. \tag{69}$$

Next, I find $\check{v}_1$ which will give the assignment between the two firms. For this I use Leibniz's rule on (65) to find the derivative with respect to $\check{v}_1$, and set the first-order derivative to zero to obtain

$$\int_{\underline{v}_2}^{\overline{v}_2} \left[\check{v}_1 q_1^\ell + v_2 q_2^\ell - C(q^\ell) \right] f(\check{v}_1) \mathrm{d}F_2 - \int_{\underline{v}_2}^{\overline{v}_2} \left[\check{v}_1 q_1^h + v_2 q_2^h - C(q^h) \right] f(\check{v}_1) \mathrm{d}F_2 = 0. \tag{70}$$

Re-arranging (70) gives

$$\int_{\underline{v}_2}^{\overline{v}_2} \left[\check{v}_1 q_1^\ell + v_2 q_2^\ell - C(q^\ell) - \check{v}_1 q_1^h - v_2 q_2^h + C(q^h) \right] f(\check{v}_1) \mathrm{d}F_2 = 0. \tag{71}$$

The equation (71) implies that when solved for $\check{v}_1$ (see (84) below), the curve determining the set of indifferent consumers in (v_1, v_2) space is a function of v_2 and the socially optimal qualities.

I simplify first-order conditions (66)-(69) next. Because C_1 do not depend on v_1, I obtain the following from (66)

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\check{v}_1} v_1 \mathrm{d}F_1 \right\} \mathrm{d}F_2 - C_1(q^\ell) \int_{\underline{v}_2}^{\overline{v}_2} F_1(\check{v}_1) \mathrm{d}F_2 = 0. \tag{72}$$

Rearranging (72) gives

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\check{v}_1} v_1 \mathrm{d}F_1 \right\} \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\check{v}_1) \mathrm{d}F_2} = C_1(q^\ell). \tag{73}$$

Because C_1 do not depend on v_1, I obtain the following from (67)

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\check{v}_1} v_2 \mathrm{d}F_1 \right\} \mathrm{d}F_2 - C_2(q^\ell) \int_{\underline{v}_2}^{\overline{v}_2} F_1(\check{v}_1) \mathrm{d}F_2 = 0. \tag{74}$$

Rearranging (74) gives

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{\check{v}_1} v_2 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\check{v}_1)\mathrm{d}F_2} = C_2(q^{\ell}). \tag{75}$$

Because C_1 do not depend on v_1, I obtain the following from (68)

$$\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\check{v}_1}^{\overline{v}_1} v_1 \mathrm{d}F_1\right\}\mathrm{d}F_2 - C_1(q^h)\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1(\check{v}_1)\right]\mathrm{d}F_2 = 0. \tag{76}$$

Rearranging (76) gives

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\check{v}_1}^{\overline{v}_1} v_1 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1(\check{v}_1)\right]\mathrm{d}F_2} = C_2(q^h). \tag{77}$$

Because C_1 do not depend on v_1, I obtain the following from (69)

$$\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\check{v}_1}^{\overline{v}_1} v_2 \mathrm{d}F_1\right\}\mathrm{d}F_2 - C_2(q^h)\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1(\check{v}_1)\right]\mathrm{d}F_2 = 0. \tag{78}$$

Rearranging (78) gives

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\check{v}_1}^{\overline{v}_1} v_2 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1(\check{v}_1)\right]\mathrm{d}F_2} = C_2(q^h). \tag{79}$$

The social-surplus maximizing qualities and the assignment of consumers across the two firms are the qualities and the allocation that satisfy (73), (75), (77), (79) simultaneously with (71). Because $\check{v}_1$ that maximizes (65) depend on v_2, v_1^* is determined simultaneously with qualities $(q^{\ell *}, q^{h*})$, and thus the first-best $(q^{\ell *}, q^{h*}, v^*)$ with $q_1^{\ell} \neq q_1^h$ will be characterized by

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{v_1^*\left(v_2;q^{\ell *},q^{h}*\right)} v_1 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(v_1^*\left(v_2;q^{\ell *},q^{h*}\right)\right)\mathrm{d}F_2} = C_1\left(q^{\ell *}\right) \tag{80}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{v_1^*\left(v_2;q^{\ell *},q^{h*}\right)} v_2 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(v_1^*\left(v_2;q^{\ell *},q^{h*}\right)\right)\mathrm{d}F_2} = C_2\left(q^{\ell *}\right) \tag{81}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{v_1^*\left(v_2;q^{\ell *},q^{h*}\right)}^{\overline{v}_1} v_1 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(v_1^*\left(v_2;q^{\ell *},q^{h*}\right)\right)\right]\mathrm{d}F_2} = C_1\left(q^{h*}\right) \tag{82}$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{v_1^*\left(v_2;q^{\ell *},q^{h*}\right)}^{\overline{v}_1} v_2 \mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(v_1^*\left(v_2;q^{\ell *},q^{h*}\right)\right)\right]\mathrm{d}F_2} = C_2\left(q^{h*}\right), \tag{83}$$

and

$$v_1^*\left(v_2;q^{\ell*},q^{h*}\right)=\frac{C\left(q^{h*}\right)-C\left(q^{\ell*}\right)}{q_1^{h*}-q_1^{\ell*}}-\frac{q_2^{h*}-q_2^{\ell*}}{q_1^{h*}-q_1^{\ell*}}v_2, \tag{84}$$

of which (84) is determined from (71) pointwise for all v_2.

Last, I discuss the conditions that should be satisfied for the first-best in (3) to have a unique maximum. The necessary condition for the $\left(q^{\ell*},q^{h*},v_1^*\right)$ to be a maximum is that the first-order partial derivatives with respect to $\left(q_1^{\ell},q_2^{\ell},q_1^{h},q_2^{h},\check{v}_1\right)$ are zero. These first-order conditions are given by (73), (75), (77), (79), and (71) above. The sufficient conditions for $\left(q_1^{\ell},q_2^{\ell},q_1^{h},q_2^{h}\right)$ to be a maximum can be studied by analyzed the symmetric Hessian and its successive principal minors (evaluated at the point in which the the first-order conditions are zero $\left(q_1^{\ell},q_2^{\ell},q_1^{h},q_2^{h}\right)$). For a more detailed description of this analysis, see Appendix 3.A.5.1.

Proof of Result 1

This proof contains three parts. First, I give the explicit assumptions for the maximization of (3) to have a well-defined interior solution. Second, I assume that the per-unit cost of quality production is additively separable ($\theta=0$). Using this assumption I prove the first part of Result 1. Third, I use a proof by contradiction to show that when the per-unit cost of quality production is not separable ($\theta\neq 0$), the second part of Result 1 follows. I use the following simplifying notation $v_1^*=v_1^*\left(v_2;q^{\ell*},q^{h*}\right)$ to make notation more clear.

Assuming the properties given in Subsections 3.2.2 and 3.2.3, and the discussion in Appendix 3.A.6.1, $(q^{\ell*},q^{h*},v_1^*)$ maximizes $W(q^{\ell},q^{h})$.

Suppose $\left(q^{\ell*},q^{h*}\right)$ that maximizes $W(q^{\ell},q^{h})$ has $q_2^{\ell*}=q_2^{h*}$. Because $q_2^{\ell*}=q_2^{h*}$ and because both providers have the same per-unit production cost of quality that is $C(q)=c(q_1)+\theta q_1q_2+c(q_2)$, v_1^* given by (8) becomes:

$$v_1^*=\frac{c(q_1^{h*})+\theta\left(q_1^{h*}q_2^{h*}-q_1^{\ell*}q_2^{\ell*}\right)-c(q_1^{\ell*})}{q_1^{h*}-q_1^{\ell*}}. \tag{85}$$

To prove the first part of Result 1, I assume that the per-unit cost of quality production is additively separable, $\theta=0$. Then (85) becomes:

$$v_1^*=\frac{c(q_1^{h*})-c(q_1^{\ell*})}{q_1^{h*}-q_1^{\ell*}}. \tag{86}$$

Consider the difference between (7) and (5) next. The difference is

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{v_1^*}^{\overline{v}_1}v_2\mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(v_1^*\right)\right]\mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{v_1^*}v_2\mathrm{d}F_1\right\}\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}F_1\left(v_1^*\right)\mathrm{d}F_2}=C_2(q^{h*})-C_2(q^{\ell*}). \tag{87}$$

The second quality attributes are equal if the difference between (7) and (5) given by (87) is zero in the first best.

Thus, I manipulate (87) by the following three steps (88)-(90). First, I use (2) together with an observation that v_1^* is now given by (86) does not depend on v_2.[10] This gives LHS of (88). Also, I use $C_2(q^{h*}) = c'(q_2^{h*}) + \theta q_1^{h*}$ and $C_2(q^{\ell*}) = c'(q_2^{\ell*}) + \theta q_1^{\ell*}$ to obtain the RHS of (88):

$$\frac{[1-F_1(v_1^*)]\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{[1-F_1(v_1^*)]\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - \frac{F_1(v_1^*)\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{F_1(v_1^*)\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} = c_2(q_2^{h*}) + \theta q_1^{h*} - c_2(q_2^{\ell*}) - \theta q_1^{\ell*}. \tag{88}$$

Second, I cancel $1 - F_1(v_1^*)$ and $F_1(v_1^*)$ from the nominator and denominator in (88). This gives the LHS of (89). I obtain the RHS of (89) by re-arranging and taking θ as a common term:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} = c_2(q_2^{h*}) - c_2(q_2^{\ell*}) + \theta(q_1^{h*} - q_1^{\ell*}) \tag{89}$$

Third, for (90) I use the assumption of the separable per-unit production of quality ($\theta = 0$) and

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} = c_2(q_2^{h*}) - c_2(q_2^{\ell*}). \tag{90}$$

Because both firms have the same per-unit production cost functions, (90) holds only if $q_2^{\ell*} = q_2^{h*}$, as was initially assumed. This completes the proof of the first statement.

For the proof for the second statement I use proof by contradiction and assume that when the per-unit production cost of quality is not separable ($\theta \neq 0$), there is differentiation in both quality attributes. I proceed by assuming the contrary, that is $q_2^{\ell*} = q_2^{h*}$. Now, with non-separable costs,

$$v_1^* = \frac{c(q_1^{h*}) - c(q_1^{\ell*})}{q_1^{h*} - q_1^{\ell*}} + \theta q_2^{h*}. \tag{91}$$

Moreover, (87) can be simplified by using similar three steps as above (88)-(89). It is

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} = c_2(q_2^{h*}) - c_2(q_2^{\ell*}) + \theta(q_1^{h*} - q_1^{\ell*}) \tag{92}$$

Because $q_1^{\ell*} < q_1^{h*}$ and $\theta \neq 0$ this leads to a contradiction, gives the second statement in Result 1, and thus completes the proof.

---

10 This results from the assumption of $q_2 = r_2$.

3.A.2 Proofs of the results in Section 3.4

Proof of Lemma 1

I begin by characterizing Firm B's best response price. Then I find Firm A's best response price. The conditions that ensure that the first-order conditions are sufficient are given in Appendix A6.

Consider

$$\widehat{p}_B = \arg\max_{p_B} \int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right)\right]\left[p_B - C\left(r\right)\right] \mathrm{d}F_2,$$

with

$$\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right) = \frac{p_B - \widehat{p}_A}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}.$$

I differentiate the profit function given by (15) with respect to p_B and set it to zero to obtain the following first-order condition

$$\begin{aligned} &\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right)\right] \mathrm{d}F_2 \\ &\quad = \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right) \mathrm{d}F_2 \left[\frac{p_B - C\left(r\right)}{r_1 - q_1}\right]. \end{aligned} \tag{93}$$

Then I rearrange (93) and obtain the equilibrium price

$$\widehat{p}_B - C\left(r\right) = \frac{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right)\right] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right) \mathrm{d}F_2} \left(r_1 - q_1\right).$$

I consider Firm A's best response price next. The best response price of Firm A is

$$\begin{aligned} \widehat{p}_A = &\arg\max_{p_A} \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\widetilde{v}_1(v_2; p_A, \widehat{p}_B, q, r)} \left[v_1 q_1 + v_2 q_2 - C\left(q\right)\right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 \\ &+ \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\widetilde{v}_1(v_2; p_A, \widehat{p}_B, q, r)}^{\overline{v}_1} \left[v_1 r_1 + v_2 r_2 - C\left(r\right)\right] \mathrm{d}F_1 \right\} \mathrm{d}F_2, \end{aligned}$$

with

$$\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) = \frac{\widehat{p}_B - p_A}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}. \tag{94}$$

Firm A chooses price p_A to maximize (14), given price p_B and the curve of indifferent consumers (94). Differentiating (14) with respect to p_A using Leibniz's rule gives

$$\begin{aligned} &\int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) r_1 + v_2 r_2 - C\left(r\right)\right] f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \frac{\partial \widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)}{\partial p_A} \mathrm{d}F_2 \\ &- \int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) q_1 + v_2 q_2 - C\left(q\right)\right] f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \frac{\partial \widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)}{\partial p_A} \mathrm{d}F_2. \end{aligned} \tag{95}$$

I substitute $\frac{\partial \widetilde{v}_1(v_2;q,r)}{\partial p_A} = -\frac{1}{r_1 - q_1}$ for (95) and set the first-order derivative to zero gives:

$$\int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) r_1 + v_2 r_2 - C\left(r\right)\right] \left(-\frac{1}{r_1 - q_1}\right) f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2$$
$$- \int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) q_1 + v_2 q_2 - C\left(q\right)\right] \left(-\frac{1}{r_1 - q_1}\right) f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2 = 0. \tag{96}$$

Then, I rearrange and take common terms from (96). This gives the following

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\left(r_1 - q_1\right) - \left[C\left(r\right) - C\left(q\right) - v_2\left(r_2 - q_2\right)\right] \right\} \left(-\frac{1}{r_1 - q_1}\right) f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2 = 0.$$

Firm A's first-order condition is

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ -\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) + \left[\frac{C\left(r\right) - C\left(q\right)}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{q_1 - r_1}\right] \right\} f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2 = 0. \tag{97}$$

Because the objective function of Firm A given by the integrand (14) does not contain p_A, Firm A's maximization problem can also be thought as it choosing the curve of indifferent consumers and then the equilibrium price, $\widehat{p}_A$, is determined as a residual from this procedure. This arises from one-to-one mapping between $\widehat{p}_A$ and $\widehat{v}_1$, and therefore Firm A's optimization problem can be defined either by Firm A choosing either p_A or $\widetilde{v}_1$. The equation (97) implies that when solved for $\widehat{v}_1$ (see (98) below), the curve determining the set of indifferent consumers in (v_1, v_2) space is a function of v_2.

Thus, Firm A chooses $\widehat{p}_A$ to satisfy the condition determining the curve of consumers who are indifferent between buying from Firm A and Firm B, which is

$$\widetilde{v}_1 q_1 + v_2 q_2 - p_A = \widetilde{v}_1 r_1 + v_2 r_2 - p_B, \tag{98}$$

such that the condition (97) holds. This gives the following condition for $\widehat{v}_1$:

$$\begin{aligned} \widehat{v}_1\left(v_2; \widehat{p}_A, \widehat{p}_B, q, r\right) &= \frac{\widehat{p}_B - \widehat{p}_A}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1} \\ &= \frac{C\left(r\right) - C\left(q\right)}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}. \end{aligned}$$

Now, $\widehat{p}_A$ and $\widehat{p}_B$ are mutual best responses, and I complete the proof.

Proof of Lemma 2

In this proof, I derive how the equilibrium intercept and the equilibrium slope change with qualities. I begin by finding the derivatives of the intercept $\alpha\left(q, r\right) = \frac{C(r) - C(q)}{r_1 - q_1}$ and the slope $\beta\left(q, r\right) = \frac{r_2 - q_2}{r_1 - q_1}$ with respect to qualities q_1, q_2, r_1, and r_2.

First, I differentiate $\alpha(q,r)$ with respect to q_1, q_2, r_1, and r_2:

$$\frac{\partial \alpha(q,r)}{\partial q_1} = \frac{1}{r_1 - q_1}\left[\alpha(q,r) - C_1(q)\right] \tag{99}$$
$$\frac{\partial \alpha(q,r)}{\partial q_2} = -\frac{1}{r_1 - q_1} C_2(q) \tag{100}$$
$$\frac{\partial \alpha(q,r)}{\partial r_1} = \frac{1}{r_1 - q_1}\left[C_1(r) - \alpha(q,r)\right] \tag{101}$$
$$\frac{\partial \alpha(q,r)}{\partial r_2} = \frac{1}{r_1 - q_1} C_2(r). \tag{102}$$

Second, I differentiate $\beta(q,r)$ with respect to q_1, q_2, r_1, and r_2:

$$\frac{\partial \beta(q,r)}{\partial q_1} = \frac{1}{r_1 - q_1}\beta(q,r) \tag{103}$$
$$\frac{\partial \beta(q,r)}{\partial q_2} = -\frac{1}{r_1 - q_1} \tag{104}$$
$$\frac{\partial \beta(q,r)}{\partial r_1} = -\frac{1}{r_1 - q_1}\beta(q,r) \tag{105}$$
$$\frac{\partial \beta(q,r)}{\partial r_2} = \frac{1}{r_1 - q_1}. \tag{106}$$

The following shows how the equilibrium intercept change with qualities:

$$\begin{aligned}
\frac{\partial \alpha(q,r)}{\partial q_1} + \frac{\partial \alpha(q,r)}{\partial r_1} &= \frac{1}{r_1 - q_1}\left[-C_1(q) + \alpha(q,r)\right] + \frac{1}{r_1 - q_1}\left[C_1(r) - \alpha(q,r)\right] \\
&= \frac{C_1(r) - C_1(q)}{r_1 - q_1} \\
\frac{\partial \alpha(q,r)}{\partial q_2} + \frac{\partial \alpha(q,r)}{\partial r_2} &= -\frac{C_2(q)}{r_1 - q_1} + \frac{C_2(r)}{r_1 - q_1} \\
&= \frac{C_2(r) - C_2(q)}{r_1 - q_1}.
\end{aligned}$$

These two equations above give the first equation in Lemma 2.
The following shows how the equilibrium slope change with qualities:

$$\begin{aligned}
\frac{\partial \beta(q,r)}{\partial q_1} + \frac{\partial \beta(q,r)}{\partial r_1} &= -\frac{1}{r_1 - q_1}\beta(q,r) + \frac{1}{r_1 - q_1}\beta(q,r) = 0 \\
\frac{\partial \beta(q,r)}{\partial q_2} + \frac{\partial \beta(q,r)}{\partial r_2} &= -\frac{1}{r_1 - q_1} + \frac{1}{r_1 - q_1} = 0.
\end{aligned}$$

These two equations above give the second equation in Lemma 2.

Proof of Lemma 3

In this proof I characterize the price reaction functions for Firm A and B. I begin by differentiating $\widehat{p}_A$ and $\widehat{p}_B$ with respect to r_1 and r_2. Then I find how the equilibrium curve of indifferent consumers changes the price-cost markup of the private firm $H(\alpha,\beta) = \frac{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha,\beta)]\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha,\beta)\mathrm{d}F_2}$, that is $\frac{\partial H}{\partial \alpha}$ and $\frac{\partial H}{\partial \beta}$. Last, I complete the price reaction function characterization by substituting $\frac{\partial H}{\partial \alpha}$ and $\frac{\partial H}{\partial \beta}$ for $\frac{\partial \widehat{p}_A}{\partial r_i}$ and $\frac{\partial \widehat{p}_B}{\partial r_i}$, $i = 1,2$ and use Lemma 2.

First, I differentiate $\widehat{p}_A$ and $\widehat{p}_B$ from equation (16) with respect to r_1 and r_2

$$\frac{\partial \widehat{p}_A}{\partial r_1} = H(\alpha,\beta) + (r_1 - q_1)\left[\frac{\partial H(\alpha,\beta)}{\partial \alpha}\frac{\partial \alpha}{\partial r_1} + \frac{\partial H(\alpha,\beta)}{\partial \beta}\frac{\partial \beta}{\partial r_1}\right] \tag{107}$$

$$\frac{\partial \widehat{p}_A}{\partial r_2} = (r_1 - q_1)\left[\frac{\partial H(\alpha,\beta)}{\partial \alpha}\frac{\partial \alpha}{\partial r_2} + \frac{\partial H(\alpha,\beta)}{\partial \beta}\frac{\partial \beta}{\partial r_2}\right] \tag{108}$$

$$\frac{\partial \widehat{p}_B}{\partial r_1} = C_1(r) + H(\alpha,\beta) + (r_1 - q_1)\left[\frac{\partial H(\alpha,\beta)}{\partial \alpha}\frac{\partial \alpha}{\partial r_1} + \frac{\partial H(\alpha,\beta)}{\partial \beta}\frac{\partial \beta}{\partial r_1}\right] \tag{109}$$

$$\frac{\partial \widehat{p}_B}{\partial r_2} = C_2(r) + (r_1 - q_1)\left[\frac{\partial H(\alpha,\beta)}{\partial \alpha}\frac{\partial \alpha}{\partial r_2} + \frac{\partial H(\alpha,\beta)}{\partial \beta}\frac{\partial \beta}{\partial r_2}\right]. \tag{110}$$

Then I find how the price-cost markup of the private firm $H(\alpha,\beta)$ when the equilibrium curve of indifferent consumers change, that is I find $\frac{\partial H(\alpha,\beta)}{\partial \alpha}$ and $\frac{\partial H(\alpha,\beta)}{\partial \beta}$.

For a generic function $f(x)$, it holds that

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln f(x) = \frac{f'(x)}{f(x)}.$$

Therefore by definition

$$f'(x) = f(x)\frac{\mathrm{d}}{\mathrm{d}x}\ln f(x). \tag{111}$$

I take a natural logarithm from $H(\alpha,\beta) = \frac{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha,\beta)]\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha,\beta)\mathrm{d}F_2}$ to obtain

$$\ln H(\alpha,\beta) = \ln\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2 - \ln\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2. \tag{112}$$

I differentiate (112) with respect to α and β to obtain

$$\frac{\partial \ln H(\alpha,\beta)}{\partial \alpha} = -\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}$$

and

$$\begin{aligned}\frac{\partial \ln H(\alpha,\beta)}{\partial \beta} &= -\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)(-v_2)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha - v_2\beta)(-v_2)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2} \\ &= \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}.\end{aligned}$$

I use (111) to obtain

$$\frac{\partial H(\alpha,\beta)}{\partial\alpha}=H(\alpha,\beta)\frac{\partial}{\partial\alpha}\ln H(\alpha,\beta)$$

and

$$\frac{\partial H(\alpha,\beta)}{\partial\beta}=H(\alpha,\beta)\frac{\partial}{\partial\beta}\ln H(\alpha,\beta).$$

Then the partial derivatives of $H(\alpha,\beta)$ with respect to α and β are

$$\begin{aligned}\frac{\partial H(\alpha,\beta)}{\partial\alpha}=&\left\{\frac{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right\}\\&\left\{-\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)f_2(v_2)\,\mathrm{d}v_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right\}\end{aligned}\tag{113}$$

$$=-1-H(\alpha,\beta)\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}.\tag{114}$$

and

$$\frac{\partial H(\alpha,\beta)}{\partial\beta}=H(\alpha,\beta)\left\{\frac{\int_{\underline{v}_2}^{\overline{v}_2}v_2f_1(\alpha-v_2\beta)f_2(v_2)\,\mathrm{d}v_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2}v_2f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right\}.\tag{115}$$

Last, to complete full characterization of the strategic effects of changes in Firm B's quality on both firm's prices, I substitute (114) and (115) for (107) and use $\frac{\partial\alpha}{\partial r_1}=\frac{1}{r_1-q_1}[C_1(r)-\alpha(q,r)]$ (which is (101) from the proof of Lemma 2), and $\frac{\partial\beta}{\partial r_1}=-\frac{1}{r_1-q_1}\beta(q,r)$ (which is (105) from the proof of Lemma 2) to obtain

$$\begin{aligned}\frac{\partial\widehat{p}_A}{\partial r_1}=&H(\alpha,\beta)+(r_1-q_1)\left[\frac{\partial H(\alpha,\beta)}{\partial\alpha}\frac{\partial\alpha}{\partial r_1}+\frac{\partial H(\alpha,\beta)}{\partial\beta}\frac{\partial\beta}{\partial r_1}\right]\\=&H(\alpha,\beta)+H(\alpha,\beta)\left\{\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right]\right.\\&[C_1(r)-\alpha(q,r)]\\&\left.-\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1'(\alpha-v_2\beta)v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right]\beta(q,r)\right\}\\=&H(\alpha,\beta)\left\{1+\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2}f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right][\alpha(q,r)-C_1(r)]\right.\\&\left.-\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2}v_2f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}+\frac{\int_{\underline{v}_2}^{\overline{v}_2}v_2f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right]\beta(q,r)\right\},\end{aligned}$$

which gives (20).

Similarly, I substitute (114) and (115) for (108) and use $\frac{\partial\alpha}{\partial r_2} = \frac{C_2(r)}{r_1-q_1}$ (which is (102) from the proof of Lemma 2) and $\frac{\partial\beta}{\partial r_2} = \frac{1}{r_1-q_1}$ (which is (106) from the proof of Lemma 2)

$$\frac{\partial\widehat{p}_A}{\partial r_2} = (r_1 - q_1)\left[\frac{\partial H(\alpha,\beta)}{\partial\alpha}\frac{\partial\alpha}{\partial r_2} + \frac{\partial H(\alpha,\beta)}{\partial\beta}\frac{\partial\beta}{\partial r_2}\right]$$

$$= H(\alpha,\beta)\left\{\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}\right] C_2(r)\right.$$
$$\left. + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}\right\},$$

which gives (21).

I substitute (114) and (115) for (109) and use $\frac{\partial\alpha}{\partial r_1} = \frac{1}{r_1-q_1}[C_1(r) - \alpha(q,r)]$ (which is (101) from the proof of Lemma 2) and $\frac{\partial\beta}{\partial r_1} = -\frac{1}{r_1-q_1}\beta(q,r)$ (which is (105) from the proof of Lemma 2):

$$\frac{\partial\widehat{p}_B}{\partial r_1} = C_1(r) + H(\alpha,\beta)\left\{1 + \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}\right]\right.$$
$$[\alpha(q,r) - C_1(r)]$$
$$\left. - \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}\right]\beta(q,r)\right\},$$

which gives (22).

Last, I substitute (114) and (115) for (110) and use $\frac{\partial\alpha}{\partial r_2} = \frac{C_2(r)}{r_1-q_1}$ (which is (102) from the proof of Lemma 2) and $\frac{\partial\beta}{\partial r_2} = \frac{1}{r_1-q_1}$ (which is (106) from the proof of Lemma 2):

$$\frac{\partial\widehat{p}_B}{\partial r_2} = C_2(r) + (r_1 - q_1)\left[\frac{\partial H(\alpha,\beta)}{\partial\alpha}\frac{\partial\alpha}{\partial r_2} + \frac{\partial H(\alpha,\beta)}{\partial\beta}\frac{\partial\beta}{\partial r_2}\right]$$

$$= C_2(r) + H(\alpha,\beta)\left\{\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha - v_2\beta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}\right] C_2(r)\right.$$
$$\left. + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\, v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\alpha - v_2\beta)]\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\alpha - v_2\beta)\, v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\alpha - v_2\beta)\,\mathrm{d}F_2}\right\}$$

which gives (23).

Proof of Proposition 1

In this proof I characterize the first-order conditions for Firm A and Firm B in addition to finding the equilibrium curve of indifferent consumers.

Given the subgame-perfect equilibrium prices $\widehat{p}$, equilibrium qualities $\widehat{q}$ and $\widehat{r}$ are mutual best responses:

$$\widehat{q} \equiv (\widehat{q}_1, \widehat{q}_2) = \underset{q=(q_1,q_2)}{\arg\max} \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})} [v_1 q_1 + v_2 q_2 - C(q)] \, \mathrm{d}F_1 \right\} \mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1} [v_1 \widehat{r}_1 + v_2 \widehat{r}_2 - C(\widehat{r})] \, \mathrm{d}F_1 \right\} \mathrm{d}F_2 \quad (116)$$

and

$$\widehat{r} \equiv (\widehat{r}_1, \widehat{r}_2) = \underset{r=(r_1,r_2)}{\arg\max} \int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] \left[\widehat{p}_B(\widehat{q}, r) - C(r)\right] \mathrm{d}F_2. \quad (117)$$

Note that

$$\widehat{v}_1(q,r) \equiv \widetilde{v}_1(v_2; \widehat{p}_A(q,r), \widehat{p}_B(q,r), q, r) = \frac{\widehat{p}_B(q,r) - \widehat{p}_A(q,r)}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}. \quad (118)$$

I begin the proof by finding Firm A's first-order conditions first. I use Leibniz's rule to obtain the first-order derivative of (24) with respect to q_1

$$\int_{\underline{v}_2}^{\overline{v}_2} \int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})} [v_1 - C_1(q)] \, \mathrm{d}F_1 \mathrm{d}F_2 + \int_{\underline{v}_2}^{\overline{v}_2} [\widehat{v}_1(v_2; q, \widehat{r}) q_1 + v_2 q_2 - C(q)] f_1(\widehat{v}_1(v_2; q, \widehat{r})) \frac{\partial \widehat{v}_1(v_2; q, \widehat{r})}{\partial q_1} \mathrm{d}F_2 - \int_{\underline{v}_2}^{\overline{v}_2} [\widehat{v}_1(v_2; q, \widehat{r}) r_1 + v_2 \widehat{r}_2 - C(\widehat{r})] f_1(\widehat{v}_1(v_2; q, \widehat{r})) \frac{\partial \widehat{v}_1(v_2; q, \widehat{r})}{\partial q_1} \mathrm{d}F_2. \quad (119)$$

I rearrange the first-order derivative in (119) to obtain

$$\int_{\underline{v}_2}^{\overline{v}_2} \int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})} [v_1 - C_1(q)] \, \mathrm{d}F_1 \mathrm{d}F_2 - \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \widehat{v}_1(v_2; q, \widehat{r}) (\widehat{r}_1 - q_1) - [C(\widehat{r}) - C(q) - v_2(\widehat{r}_2 - q_2)] \right\} f_1(\widehat{v}_1(v_2; q, \widehat{r})) \frac{\partial \widehat{v}_1(v_2; q, \widehat{r})}{\partial q_1} \mathrm{d}F_2. \quad (120)$$

I set (120) equal to zero and evaluate (120) at the equilibrium qualities. By Lemma 1 the term inside the curly brackets of (120) is zero. This gives the following first-order condition:

$$\int_{\underline{v}_2}^{\overline{v}_2} \int_{\underline{v}_1}^{\widehat{v}_1(v_2;\widehat{q},\widehat{r})} [v_1 - C_1(\widehat{q})] \, \mathrm{d}F_1 \mathrm{d}F_2 = 0. \quad (121)$$

Lastly, I rearrange (121), which gives:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \int_{\underline{v}_1}^{\widehat{v}_1(v_2;\widehat{q},\widehat{r})} v_1 \mathrm{d}F_1 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2; \widehat{q}, \widehat{r})) \, \mathrm{d}F_2} = C_1(\widehat{q}). \quad (122)$$

Next, I derive the first-order condition for Firm A with respect to q_2. I obtain the first-order derivative of (24) with respect to q_2 by using Leibniz's rule

$$\begin{aligned}\int_{\underline{v}_2}^{\overline{v}_2}\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})}&[v_2 - C_2(q)]\,\mathrm{d}F_1\mathrm{d}F_2\\ &+\int_{\underline{v}_2}^{\overline{v}_2}[\widehat{v}_1(v_2;q,\widehat{r})\,q_1 + v_2q_2 - C(q)]\,f_1(\widehat{v}_1(v_2;q,\widehat{r}))\,\frac{\partial\widehat{v}_1(v_2;q,\widehat{r})}{\partial q_2}\mathrm{d}F_2\\ -&\int_{\underline{v}_2}^{\overline{v}_2}[\widehat{v}_1(v_2;q,\widehat{r})\,\widehat{r}_1 + v_2\widehat{r}_2 - C(\widehat{r})]\,f_1(\widehat{v}_1(v_2;q,\widehat{r}))\,\frac{\partial\widehat{v}_1(v_2;q,\widehat{r})}{\partial q_2}\mathrm{d}F_2.\end{aligned} \tag{123}$$

I rearrange (123) to obtain

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2}\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})}[v_2 - C_2(q)]\,\mathrm{d}F_1\mathrm{d}F_2\\ +&\int_{\underline{v}_2}^{\overline{v}_2}\left\{\widehat{v}_1(v_2;q,\widehat{r})\,(\widehat{r}_1 - q_1) - [C(\widehat{r}) - C(q) - v_2(\widehat{r}_2 - q_2)]\right\}\\ &\quad f_1(\widehat{v}_1(v_2;q,\widehat{r}))\,\frac{\partial\widehat{v}_1(v_2;q,\widehat{r})}{\partial q_2}\mathrm{d}F_2.\end{aligned} \tag{124}$$

I set (124) equal to zero and evaluate (124) at the equilibrium qualities. By Lemma 1 the equilibrium the term inside the curly brackets of (124) is zero. This gives the following first-order condition:

$$\int_{\underline{v}_2}^{\overline{v}_2}\int_{\underline{v}_1}^{\widehat{v}_1(v_2;\widehat{q},\widehat{r})}[v_2 - C_2(\widehat{q})]\,\mathrm{d}F_1\mathrm{d}F_2 = 0. \tag{125}$$

Lastly, I rearrange (125) which gives:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\int_{\underline{v}_1}^{\widehat{v}_1(v_2;\widehat{q},\widehat{r})} v_2\mathrm{d}F_1\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2;\widehat{q},\widehat{r}))\,\mathrm{d}F_2} = C_2(\widehat{q})\,. \tag{126}$$

Differentiating Firm B's profits with respect to r_1 are

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\widehat{v}_1(v_2;\widehat{q},r))]\,[-C_1(r)]\,\mathrm{d}F_2\\ +&\frac{\partial}{\partial r_1}\left\{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\widehat{v}_1(v_2;\widehat{q},r))]\,\mathrm{d}F_2\right\}\times[\widehat{p}_B(\widehat{q},r) - C(r)] \qquad (127)\\ +&\frac{\partial}{\partial\widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2}[1 - F_1(\widehat{v}_1(v_2;\widehat{q},r))]\,\mathrm{d}F_2\right\}\frac{\partial\widehat{p}_A}{\partial r_1}\times[\widehat{p}_B(\widehat{q},r) - C(r)]\,, \qquad (128)\end{aligned}$$

of which the effect via Firm B's own equilibrium price has been ignored due to Lemma 1 (the effect via $\widehat{p}_B$ on Firm B's own price is second order due to the Envelope theorem). The first two rows in (127) give how private firm's quality affects its cost and demand. The last row in (128) describes the strategic effect of

Firm B's quality on Firm A's price. To obtain the complete partial derivative, I use $\frac{\partial \widehat{v}_1}{\partial r_1} = -\frac{\widehat{v}_1}{r_1 - q_1}$ to obtain[11]

$$\frac{\partial}{\partial r_1}\left\{ \int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; \widehat{q}, r\right)\right)\right] \mathrm{d}F_2 \right\} = \frac{1}{r_1 - q_1} \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, r\right)\right) \widehat{v}_1\left(v_2; \widehat{q}, r\right) \mathrm{d}F_2 \tag{129}$$

and I use $\frac{\partial \widehat{v}_1}{\partial p^A} = -\frac{1}{r_1 - q_1}$ to obtain

$$\frac{\partial}{\partial \widehat{p}_A}\left\{ \int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; \widehat{q}, r\right)\right)\right] \mathrm{d}F_2 \right\} = \frac{1}{r_1 - q_1} \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, r\right)\right) \mathrm{d}F_2. \tag{130}$$

I use (129) and (130) for (127) and (128) to obtain

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; \widehat{q}, r\right)\right)\right] \left[-C_1(r)\right] \mathrm{d}F_2 \\
&\quad + \underbrace{\frac{1}{r_1 - q_1} \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, r\right)\right) \widehat{v}_1\left(v_2; \widehat{q}, r\right) \mathrm{d}F_2}_{= \frac{\partial}{\partial r_1}\left\{ \int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] \mathrm{d}F_2 \right\}} \left[\widehat{p}_B\left(\widehat{q}, r\right) - C(r)\right] \\
&\quad + \underbrace{\frac{1}{r_1 - q_1} \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, r\right) \mathrm{d}F_2\right.}_{= \frac{\partial}{\partial \widehat{p}_A}\left\{ \int_{\underline{v}_2}^{\overline{v}_2} [1 - F_1(\widehat{v}_1(v_2; \widehat{q}, r))] \mathrm{d}F_2 \right\}} \frac{\partial \widehat{p}_A}{\partial r_1} \left[\widehat{p}_B\left(\widehat{q}, r\right) - C(r)\right].
\end{aligned} \tag{131}$$

I set the first-order derivative in (131) equal to zero and evaluate it at equilibrium qualities. Because in price equilibrium

$$\frac{\widehat{p}_B - C(r)}{r_1 - q_1} = \frac{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; q, r\right)\right)\right] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; q, r\right)\right) \mathrm{d}F_2}, \tag{132}$$

I can simplify the first-order condition (131) as follows

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right)\right] \left[-C_1(\widehat{r})\right] \mathrm{d}F_2 \\
&\quad + \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right) \widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right) \mathrm{d}F_2 \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right)\right] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right) \mathrm{d}F_2}\right] \\
&\quad + \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right) \mathrm{d}F_2 \frac{\partial \widehat{p}_A}{\partial r_1} \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right)\right] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right) \mathrm{d}F_2}\right] = 0.
\end{aligned}$$

Taking common terms and subtracting gives

$$\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right)\right] \mathrm{d}F_2 \left[\frac{\partial \widehat{p}_A}{\partial r_1} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right) \widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right) \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2; \widehat{q}, \widehat{r}\right)\right) f_2\left(v_2\right) \mathrm{d}v_2} - C_1(\widehat{r})\right] = 0. \tag{133}$$

---

11 Using (13) to gives $\frac{\partial \widehat{v}_1}{\partial r_1} = \frac{-(\widehat{p}_B - \widehat{p}_A) + v_2(r_2 - q_2)}{(r_1 - q_1)^2} = -\frac{1}{r_1 - q_1}\left[\frac{\widehat{p}_B - \widehat{p}_A}{r_1 - q_1} - \frac{v_2(r_2 - q_2)}{r_1 - q_1}\right]$.

Thus, the first-order condition (133) is

$$\frac{\partial \widehat{p}_A}{\partial r_1}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\beta\right)\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right)f_2\left(v_2\right)\mathrm{d}v_2}=C_1\left(\widehat{r}\right), \tag{134}$$

from where I have omitted the factor $\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right)\right]$.

I derive Firm B's first-order condition with respect to r_2 next. I differentiate Firm B's profits (25) with respect to r_2

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\left[C_2\left(r\right)\right]\mathrm{d}F_2\\&+\frac{\partial}{\partial r_2}\left\{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\mathrm{d}F_2\right\}\times\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right] &(135)\\&+\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\mathrm{d}F_2\right\}\frac{\partial \widehat{p}_A}{\partial r_2}\times\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right], &(136)\end{aligned}$$

from which the effect on Firm B's price via its own equilibrium price has been ignored due to Lemma 1 (the effect via $\widehat{p}_B$ on firms' own price is second order). To obtain the complete partial derivative, I use $\frac{\partial \widehat{v}_1}{\partial r_2}=-\frac{v_2}{r_1-q_1}$ to obtain

$$\frac{\partial}{\partial r_2}\left\{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\mathrm{d}F_2\right\}=\frac{1}{r_1-q_1}\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)v_2\mathrm{d}F_2 \tag{137}$$

and use $\frac{\partial \widehat{v}_1}{\partial p^A}=-\frac{1}{r_1-q_1}$ to obtain

$$\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\mathrm{d}F_2\right\}=\frac{1}{r_1-q_1}\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2. \tag{138}$$

I use (137) and (138) for (135) and (136) to obtain

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]f_2\left(v_2\right)\left[-C_2\left(r\right)\right]\mathrm{d}F_2\\&\quad+\underbrace{\frac{1}{r_1-q_1}\int_{\underline{v}_2}^{\overline{v}_2}\left[f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]v_2\mathrm{d}F_2}_{=\frac{\partial}{\partial r_2}\left\{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\mathrm{d}F_2\right\}}\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right]\\&\quad+\underbrace{\frac{1}{r_1-q_1}\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2}_{=\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\mathrm{d}F_2\right\}}\frac{\partial \widehat{p}_A}{\partial r_2}\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right]. &(139)\end{aligned}$$

I set the first-order derivative in (139) equal to zero and evaluate it at equilibrium qualities. Because in price equilibrium

$$\frac{\widehat{p}_B-C\left(r\right)}{r_1-q_1}=\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left[1-F_1\left(\widehat{v}_1\left(v_2;q,r\right)\right)\right]\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;q,r\right)\right)\mathrm{d}F_2}, \tag{140}$$

I can simplify the first-order condition as follows

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right)\right] f_2\left(v_2\right)\left[-C_2\left(\widehat{r}\right)\right] \mathrm{d}F_2 \\
&+ \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) v_2 \mathrm{d}F_2 \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right)\right] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\, \widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) \mathrm{d}F_2}\right] \\
&+ \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) \mathrm{d}F_2 \frac{\partial \widehat{p}_A}{\partial r_2}\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right)\right] \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) \mathrm{d}F_2}\right] = 0,
\end{aligned}$$

which after re-arranging becomes

$$\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right)\right] \mathrm{d}F_2 \left[\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) \mathrm{d}F_2} - C_2\left(\widehat{r}\right)\right] = 0. \tag{141}$$

Thus, the first-order condition (141) is

$$\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right) \mathrm{d}F_2} = C_2\left(\widehat{r}\right), \tag{142}$$

from where I have omitted the factor $\left[1 - F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]$.

Last, I consider the equilibrium curve of indifferent consumers. Because equilibrium prices $\widehat{p}_A$ and $\widehat{p}_B$ have to follow Lemma 1,

$$\widehat{v}_1\left(v_2;q,r\right) = \frac{C\left(r\right) - C\left(q\right)}{r_1 - q_1} - v_2 \frac{r_2 - q_2}{r_1 - q_1}$$

This together with equations (122), (126), (134), and (142) give (31)-(34) and (35) in Proposition 1, and completes the proof.

Proof of Result 2

This proof contains three parts. First, I make explicit the assumptions required for the maximization of (24) and (25) to have a well-defined interior solution. Second, I assume that the per-unit cost of quality production is separable ($\theta = 0$) and prove the first part of Result 2. Third, I use a proof by contradiction to show that when the per-unit cost of quality production is not separable ($\theta \neq 0$), the second part of Result 2 follows. A discussion of the conditions that ensure the existence of an equilibrium are given in Appendix 3.A.5.2.

I begin the proof by simplifying the curve of indifferent consumers as follows. To prove the first part of Result 2, suppose the $\widehat{q}$ that maximizes social surplus and the $\widehat{r}$ that maximizes profits have $\widehat{q}_2 = \widehat{r}_2$. Because $\widehat{q}_2 = \widehat{r}_2$, both firms have the same per-unit production cost of quality, the per-unit production cost of quality is $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$, and $\theta = 0$, $\widehat{v}_1$ given by (35) becomes:

$$\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right) = \frac{c(r_1) + \theta\left(r_1 r_2 - q_1 q_2\right) - c(q_1)}{r_1 - q_1}. \tag{143}$$

Note that $\widehat{v}_1$ given by (86) does not depend on v_2.

The difference between Firm B's first-order condition with respect to r_2 (34) and Firm A's first-order condition with respect to q_2 (32) is:

$$\left[\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha - v_2\beta\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha - v_2\beta\right) \mathrm{d}F_2}\right] - \frac{\int_{\underline{v}_2}^{\overline{v}_2}\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,r)} v_2 \mathrm{d}F_1 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\alpha - v_2\beta\right)\mathrm{d}F_2} = C_2\left(r\right) - C_2\left(q\right) \tag{144}$$

The second quality attributes are equal if the difference between (32) and (34) is zero, so that the equation (144) holds.

I begin by simplifying Firm B's first-order condition. When $\mathrm{d}F_i$ is a step function, the derivative of the valuation density is zero. Using this, and using (2), the price response function $\frac{\partial \widehat{p}_A}{\partial r_2}$ in Lemma 1 simplify as follows:

$$\begin{aligned}\frac{\partial \widehat{p}_A}{\partial r_2} &= H\left(\widehat{v}_1\right)\left\{\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\right)\right]\mathrm{d}F_2}\right] C_2\left(\,r\right) + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\right)\right]\mathrm{d}F_2}\right\} \\ &= -\,C_2\left(r\right) + H\left(\widehat{v}_1\right)\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\right)\right]\mathrm{d}F_2}.\end{aligned} \tag{145}$$

Substituting (145) for $\frac{\partial \widehat{p}_A}{\partial r_2}$ into Firm B's first-order condition with respect to r_2, that is (34)

$$-C_2\left(r\right) + H\left(\widehat{v}_1\right)\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\right)\right]\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right)\mathrm{d}F_2} - C_2\left(r\right) = 0 \tag{146}$$

Because $\widehat{v}_1$ given by (143) does not depend on v_2, $f_1\left(\widehat{v}_1\right)$ and $1 - F_1\left(\widehat{v}_1\right)$ can be taken out of the integrals. Doing this and re-arranging (146) gives

$$-2C_2\left(r\right) + \underbrace{H\left(\widehat{v}_1\right)}_{=\frac{1-F(\widehat{v}_1)}{f(\widehat{v}_1)}}\frac{f_1\left(\widehat{v}_1\right)}{1 - F_1\left(\widehat{v}_1\right)}\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\mathrm{d}F_2} + \frac{f_1\left(\widehat{v}_1\right)}{f_1\left(\widehat{v}_1\right)}\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\mathrm{d}F_2} = 0 \tag{147}$$

and after canceling the common terms gives

$$-2C_2\left(r\right) + 2\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\mathrm{d}F_2} = 0 \tag{148}$$

which becomes

$$-C_2\left(r\right)+\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}=0. \tag{149}$$

Then I simplify Firm A's first-order condition with respect to r_2. Again, because $\mathrm{d}F_i$ is a step function, the derivative of the valuation density is zero, using (2), and because $\widehat{v}_1$ given by (143) does not depend on v_2, $f_1\left(\widehat{v}_1\right)$ and $1-F_1\left(\widehat{v}_1\right)$ can be taken out of the integrals:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}-C_2(q)=0. \tag{150}$$

Combining (149) and (150) gives

$$-C_2\left(r\right)+\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}+C_2\left(q\right)=0$$

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}=C_2\left(r\right)-C_2\left(q\right). \tag{151}$$

Because $C(q)=c\left(q_1\right)+\theta q_1 q_2+c\left(q\right)$ and $\theta=0$, (151) becomes

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}=c_2\left(\widehat{r}_2\right)-c_2\left(\widehat{q}_2\right). \tag{152}$$

Because both firms have the same per-unit production cost functions, and because the LHS of (152) is zero, equation (152) can hold only if $\widehat{q}_2=\widehat{r}_2$. This completes the proof of the first statement.

For the second statement I assume the contrary, that when the per-unit production cost of quality is not separable $(\theta \neq 0)$, there is differentiation in both attributes. I proceed by assuming the contrary, that is $\widehat{q}_2=\widehat{r}_2$. Now, with non-separable costs and because both firms have the same technology, (151) becomes:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}-\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2}=c_2(r_2)-c_2(q_2)+\theta\left(r_1-q_1\right) \tag{153}$$

Because $r_1>q_1$ and $\theta \neq 0$, this leads to a contradiction, gives the second statement in Result 2, and completes the proof.

Proof of Result 3

Note that in the equilibrium characterization, Firm A's equations (equations (31) and (32)) are the same as the ones in the first best characterization (equations (4) and (5)). Also, functions that characterize the equilibrium curve of indifferent consumers and the curve of indifferent consumers in the first best have the same form (equations (35) and (8) respectively).

In this proof I compare Firm B's first-order conditions with respect to both quality attributes with the corresponding equations in the first best characterization. The equilibrium allocation is efficient if the form of the equations are the same after I have imposed the assumptions of f_i being uniform and the per-unit production costs. For the first quality attribute characterizations (the first best and equilibrium) I also use the results from Laine and Ma (2017).

I begin the proof by simplifying the curve of indifferent consumers as follows. Suppose the $\widehat{q}$ that maximizes the payoffs of Firm A and the $\widehat{r}$ that maximizes the payoffs of Firm B have $\widehat{q}_2 = \widehat{r}_2$. Because $\widehat{q}_2 = \widehat{r}_2$, both firms have the same per-unit production cost of quality, the per-unit production cost of quality has the form of $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$, and $\theta = 0$, hence $\widehat{v}_1$ given by (35) becomes:

$$\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right) = \frac{c\left(\widehat{r}_1\right) - c\left(\widehat{q}_1\right)}{\widehat{r}_1 - \widehat{q}_1}. \tag{154}$$

Note that now $\widehat{v}_1$ given by (154) does not depend on v_2. Moreover, using the same assumptions on the quality-valuation distributions and the per-unit production cost, the curve of indifferent consumers in the first best is

$$v_1^*\left(v_2;q^{\ell *},q^{h*}\right) = \frac{c(q_1^{h*}) - c(q_1^{\ell *})}{q_1^{h*} - q_1^{\ell *}}. \tag{155}$$

I continue with the proof by simplifying the equations for the second quality attributes next. Consider equation (7) from the first best characterization. Using (2), because (154) does not depend on v_2, and the assumption on the per-unit production cost, equation (7) simplifies as follows:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - c_2(q_2^h) = 0. \tag{156}$$

Then I simplify the following first-order condition:

$$\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha - v_2\beta\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\alpha - v_2\beta\right) \mathrm{d}F_2} - C(r) = 0. \tag{157}$$

Following exactly the same steps that give the equations (145)-(149) from the proof of Result 2, the first-order condition becomes

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - c_2(r_2) = 0. \tag{158}$$

Equations (156) and (158) have the same form. I move on to analyze if the equations for the first quality attributes, i.e. the equations that characterize the first best and equilibrium, have the same form.

First, I substitute $\widehat{v}_1$ for $H(\alpha, \beta)$, use the result of $\widehat{v}_1$ being independent of v_2, and use the assumption of f_2 being uniform. Thus H becomes $H(\widehat{v}_1) = \frac{1-F(\widehat{v}_1)}{f(\widehat{v}_1)}$.

Second, I simplify the first-order condition with respect to q_1^h. I use (2), because (154) does not depend on v_2, and I use the assumption on the per-unit production cost. Thus, the first-order condition with respect to q_1^h, that is the equation (6), simplifies as follows:

$$\frac{\int_{\widehat{v}_1}^{\overline{v}_1} v_1 \mathrm{d}F_1}{1 - F(\widehat{v}_1)} - c_1(q^h). \tag{159}$$

Using Lemma 3, from Laine and Ma (2017), I can rewrite (159) as

$$\widehat{v}_1 + \frac{\int_{\widehat{v}_1}^{\overline{v}_1} f(v)H(v)\mathrm{d}v_1}{1 - F(\widehat{v}_1)H(\widehat{v}_1)} - c_1(q^h), \tag{160}$$

where $\widehat{v}_1$ is given by (155).

Third, I simplify Firm B's first-order condition with respect to r_1 that is given by equation (33) next. Because f_i is uniform, $\widehat{v}_1$ is given by (154), and because $\theta = 0$ the per-unit production costs are $C(q) = c(q_1) + c(q_2)$, Firm B's price response function becomes

$$\frac{\partial \widehat{p}_A}{\partial r_1} = H\left(\widehat{v}_1\right) \left\{ 1 + \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right) \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\widehat{v}_1\right)\right] \mathrm{d}F_2} \right] \left[\widehat{v}_1 - c_1\left(r_1\right)\right] \right\} \tag{161}$$

$$= H\left(\widehat{v}_1\right) + \widehat{v}_1 - c_1\left(r_1\right), \tag{162}$$

in which I have used the assumption of $\widehat{v}_1$ being independent of v_2 and f_2 being uniform to move from step (161) to (162). Then substituting (162) for (33) and using (again) the assumptions of uniform f_1 and $\widehat{v}_1$ being independent of v_2, I can simplify the first-order condition as follows

$$\widehat{v}_1 + \frac{H\left(\widehat{v}_1\right)}{2} - c_1\left(r\right) = 0. \tag{163}$$

Because $H(\widehat{v}_1) = \frac{1-F(\widehat{v}_1)}{f(\widehat{v}_1)}$ and because f is uniform $H'(\widehat{v}_1) = -1$.

Now, equation (163) has exactly the same form as the LHS of the equation (20) in Laine and Ma (2017). Also, equation (160) has exactly the same form as the RHS of equation (20) in Laine and Ma (2017). Last, for uniform f_i the inverse hazard rate is linear, so I can use Proposition 2 from Laine and Ma (2017), which says the equilibrium qualities q_1, r_1, and market shares are first best. Then because also the equations (156) and (158) have the same form and all equations have to hold together (equilibrium condition), equilibrium qualities $\widehat{q}_1$, $\widehat{r}_1$, and $\widehat{r}_2 = \widehat{q}_2$ are the first best. This completes the proof.

3.A.3 Proofs of the results in Section 3.5

Proof of Lemma 4

I begin by characterizing Firm B's best response price. Then I find Firm A's best response price. A discussion of the conditions that ensure the existence of an equilibrium are given in Appendix A6.

The best response price of Firm B is

$$\widehat{p}_B = \arg\max_{p_B} \int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right)\left[p_B - C\left(r\right)\right] \mathrm{d}F_2,$$

in which

$$\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right) = \frac{\widehat{p}_A - p_B}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}.$$

I differentiate the profit function given by (46) with respect to p_B and set it to zero. This gives the following first-order condition:

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right) \mathrm{d}F_2 \\ &= \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right) \mathrm{d}F_2 \left[\frac{p_B - C\left(r\right)}{q_1 - r_1}\right]. \end{aligned} \tag{164}$$

Then I rearrange (164) and obtain the following equilibrium price

$$\widehat{p}_B - C\left(r\right) = \frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right) \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widetilde{v}_1\left(v_2; \widehat{p}_A, p_B, q, r\right)\right) \mathrm{d}F_2}\left(q_1 - r_1\right).$$

I consider Firm A's best response price next. The best response price of Firm A is

$$\begin{aligned}\widehat{p}_A = &\arg\max_{p_A} \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\underline{v}_1}^{\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)} \left[v_1 r_1 + v_2 r_2 - C\left(r\right)\right] \mathrm{d}F_1 \right\} \mathrm{d}F_2 \\ &+ \int_{\underline{v}_2}^{\overline{v}_2} \left\{ \int_{\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)}^{\overline{v}_1} \left[v_1 q_1 + v_2 q_2 - C\left(q\right)\right] \mathrm{d}F_1 \right\} \mathrm{d}F_2,\end{aligned}$$

with

$$\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) = \frac{p_A - \widehat{p}_B}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}. \tag{165}$$

Firm A chooses price p_A to maximize (45), given price p_B and the curve of indifferent consumers (44). Differentiating (45) with respect to p_A using Leibniz's rule gives

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) r_1 + v_2 r_2 - C\left(r\right)\right] f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \frac{\partial \widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)}{\partial p_A} \mathrm{d}F_2 \\ &- \int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) q_1 + v_2 q_2 - C\left(q\right)\right] f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \frac{\partial \widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)}{\partial p_A} \mathrm{d}F_2.\end{aligned} \tag{166}$$

I substitute $\frac{\partial \widetilde{v}_1(v_2;q,r)}{\partial p_A} = \frac{1}{q_1 - r_1}$ for (166) and set the first-order derivative to zero to obtain:

$$\int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) r_1 + v_2 r_2 - C\left(r\right)\right] \left(\frac{1}{q_1 - r_1}\right) f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2$$
$$- \int_{\underline{v}_2}^{\overline{v}_2} \left[\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) q_1 + v_2 q_2 - C\left(q\right)\right] \left(\frac{1}{q_1 - r_1}\right) f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2 = 0. \tag{167}$$

Last, I rearrange and take common terms from (167). This gives the following

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\left(q_1 - r_1\right) - \left[C\left(q\right) - C\left(r\right) - v_2\left(q_2 - r_2\right)\right] \right\} \left(\frac{1}{q_1 - r_1}\right) f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2 = 0.$$

Firm A's first-order condition is

$$\int_{\underline{v}_2}^{\overline{v}_2} \left\{ \widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right) - \left[\frac{C\left(q\right) - C\left(r\right)}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}\right] \right\} f_1\left(\widetilde{v}_1\left(v_2; p_A, \widehat{p}_B, q, r\right)\right) \mathrm{d}F_2 = 0. \tag{168}$$

Because the objective function of Firm A given by the integrand (45) does not contain p_A, Firm A's maximization problem can also be thought as it choosing the curve of indifferent consumers and then the equilibrium price, $\widehat{p}_A$, is determined as a residual from this procedure. This arises from one-to-one mapping between $\widehat{p}_A$ and $\widehat{v}_1$, and therefore Firm A's optimization problem can be defined either by Firm A choosing either p_A or $\widetilde{v}_1$. The equation (168) implies that when solved for $\widehat{v}_1$ (see (98) below), the curve determining the set of indifferent consumers in (v_1, v_2) space is a function of v_2.

Thus, Firm A chooses $\widehat{p}_A$ to satisfy the condition determining the curve of consumers who are indifferent between buying from Firm A and Firm B, which is

$$\widetilde{v}_1 q_1 + v_2 q_2 - p_A = \widetilde{v}_1 r_1 + v_2 r_2 - p_B, \tag{169}$$

such that the condition (168) holds. This gives the following condition for $\widehat{v}_1$:

$$\begin{aligned} \widehat{v}_1\left(v_2; \widehat{p}_A, \widehat{p}_B, q, r\right) &= \frac{\widehat{p}_A - \widehat{p}_B}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1} \\ &= \frac{C\left(q\right) - C\left(r\right)}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}. \end{aligned}$$

Now, $\widehat{p}_A$ and $\widehat{p}_B$ are mutual best responses, and I complete the proof.

Additional result

The following lemma describes how the equilibrium intercept and the equilibrium slope of $\widehat{v}_1$ change with qualities. As the results are similar to the ones in Section 3.4, I provide this in the Appendix instead.

Lemma A *In any subgame (q, r), in price equilibrium:*

$$\frac{\partial \gamma (q,r)}{\partial q_i} + \frac{\partial \gamma (q,r)}{\partial r_i} = \frac{C_i (q) - C_i (r)}{q_1 - r_1} \textit{ and } \frac{\partial \delta (q,r)}{\partial q_i} + \frac{\partial \delta (q,r)}{\partial r_i} = 0$$

for $i = 1, 2$.
If a firm changes either of its quality attributes, this change will affect the equilibrium prices of both firms and the equilibrium curve of indifferent consumers. Lemma A says that changes in qualities will have different effects on the intercept $\delta (q,r)$ and the slope $\gamma (q,r)$. The change in the intercept is $\frac{C_i(q)-C_i(r)}{q_1-r_1}$. The effects of changes in firms' qualities on the slope of the equilibrium set of indifferent consumers are equal and opposite.

Proof of Lemma A

In this proof, I show how the equilibrium intercept and the equilibrium slope change with qualities. First, I find the derivatives of the intercept $\gamma(q,r)$ and the slope $\delta(q,r)$ with respect to qualities q_1, q_2, r_1, and r_2. I use some parts of the proof also in the proofs below.

First, I differentiate $\gamma(q,r)$ with respect to q_1, q_2, r_1, and r_2:

$$\frac{\partial \gamma (q,r)}{\partial q_1} = \frac{1}{q_1 - r_1} \left[C_1 (q) - \gamma (q,r) \right] \tag{170}$$
$$\frac{\partial \gamma (q,r)}{\partial q_2} = \frac{1}{q_1 - r_1} C_2 (q) \tag{171}$$
$$\frac{\partial \gamma (q,r)}{\partial r_1} = \frac{1}{q_1 - r_1} \left[\gamma (q,r) - C_1 (r) \right] \tag{172}$$
$$\frac{\partial \gamma (q,r)}{\partial r_2} = -\frac{1}{q_1 - r_1} C_2 (r) \, . \tag{173}$$

Second, I differentiate $\delta(q,r)$ (section 3.5.1.) with respect to q_1, q_2, r_1, and r_2:

$$\frac{\partial \delta (q,r)}{\partial q_1} = -\frac{1}{q_1 - r_1} \delta (q,r) \tag{174}$$
$$\frac{\partial \delta (q,r)}{\partial q_2} = \frac{1}{q_1 - r_1} \tag{175}$$
$$\frac{\partial \delta (q,r)}{\partial r_1} = \frac{1}{q_1 - r_1} \delta (q,r) \tag{176}$$
$$\frac{\partial \delta (q,r)}{\partial r_2} = -\frac{1}{q_1 - r_1} . \tag{177}$$

The following shows how the equilibrium intercept changes with qualities

$$\begin{aligned}\frac{\partial\gamma(q,r)}{\partial q_1}+\frac{\partial\gamma(q,r)}{\partial r_1}&=\frac{1}{q_1-r_1}\left[C_1(q)-\gamma(q,r)\right]+\frac{1}{q_1-r_1}\left[\gamma(q,r)-C_1(r)\right]\\&=\frac{C_1(q)-C_1(r)}{q_1-r_1}\\\frac{\partial\gamma(q,r)}{\partial q_2}+\frac{\partial\gamma(q,r)}{\partial r_2}&=\frac{C_2(q)}{q_1-r_1}-\frac{C_2(r)}{q_1-r_1}\\&=\frac{C_2(q)-C_2(r)}{q_1-r_1}.\end{aligned}$$

These two equations above give the first equation in Lemma A.

The following shows how the equilibrium slope changes with qualities

$$\begin{aligned}\frac{\partial\delta(q,r)}{\partial q_1}+\frac{\partial\delta(q,r)}{\partial r_1}&=-\frac{1}{q_1-r_1}\delta(q,r)+\frac{1}{q_1-r_1}\delta(q,r)=0,\\\frac{\partial\delta(q,r)}{\partial q_2}+\frac{\partial\delta(q,r)}{\partial r_2}&=-\frac{1}{q_1-r_1}+\frac{1}{q_1-r_1}=0.\end{aligned}$$

These two equations above give the second equation in Lemma A.

Proof of Lemma 5

In this proof I characterize the price reaction functions for Firm A and B. I begin by differentiating $\widehat{p}_A$ and $\widehat{p}_B$ with respect to r_1 and r_2. Then I show how the equilibrium curve of indifferent consumers change the price-cost markup of Firm B, $G(\gamma,\delta)\equiv\frac{\int_{\underline{v}_2}^{\overline{v}_2}F_1(\gamma-v_2\delta)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}f_1(\gamma-v_2\delta)\mathrm{d}F_2}$. For this, I find the derivatives $\frac{\partial G}{\partial\gamma}$ and $\frac{\partial G}{\partial\delta}$. I complete the price reaction function characterization by substituting $\frac{\partial G}{\partial\gamma}$ and $\frac{\partial G}{\partial\delta}$ into $\frac{\partial\widehat{p}_A}{\partial r_1}$ and $\frac{\partial\widehat{p}_A}{\partial r_2}$ and use the results in the proof of Lemma A.

First, I differentiate $\widehat{p}_A$ and $\widehat{p}_B$ from Lemma 4 with respect to r_1 and r_2 to obtain

$$\frac{\partial\widehat{p}_A}{\partial r_1}=-G(\gamma,\delta)+(q_1-r_1)\left[\frac{\partial G(\gamma,\delta)}{\partial\gamma}\frac{\partial\gamma}{\partial r_1}+\frac{\partial G(\gamma,\delta)}{\partial\delta}\frac{\partial\delta}{\partial r_1}\right]\tag{178}$$

$$\frac{\partial\widehat{p}_A}{\partial r_2}=(q_1-r_1)\left[\frac{\partial G(\gamma,\delta)}{\partial\gamma}\frac{\partial\gamma}{\partial r_2}+\frac{\partial G(\gamma,\delta)}{\partial\delta}\frac{\partial\delta}{\partial r_2}\right]\tag{179}$$

$$\frac{\partial\widehat{p}_B}{\partial r_1}=C_1(r)-G(\gamma,\delta)+(q_1-r_1)\left[\frac{\partial G(\gamma,\delta)}{\partial\gamma}\frac{\partial\gamma}{\partial r_1}+\frac{\partial G(\gamma,\delta)}{\partial\delta}\frac{\partial\delta}{\partial r_1}\right]\tag{180}$$

$$\frac{\partial\widehat{p}_B}{\partial r_2}=C_2(r)+(q_1-r_1)\left[\frac{\partial G(\gamma,\delta)}{\partial\gamma}\frac{\partial\gamma}{\partial r_2}+\frac{\partial G(\gamma,\delta)}{\partial\delta}\frac{\partial\delta}{\partial r_2}\right].\tag{181}$$

Then I find how changes in the intercept and the slope of the equilibrium curve of indifferent consumers change the price-cost markup of the private firm $G(\delta,\gamma)$, that is I find $\frac{\partial G(\gamma,\delta)}{\partial\alpha}$ and $\frac{\partial G(\gamma,\delta)}{\partial\delta}$.

For a generic function $f(x)$, it holds that

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln f(x) = \frac{f'(x)}{f(x)}.$$

Thus, by definition

$$f'(x) = f(x)\frac{\mathrm{d}}{\mathrm{d}x}\ln f(x). \tag{182}$$

I take a natural logarithm from $G(\gamma, \delta)$ to obtain

$$\ln G(\gamma, \delta) = \ln \int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2 - \ln \int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2. \tag{183}$$

I differentiate (183) with respect to γ and δ to obtain

$$\frac{\partial \ln G(\gamma, \delta)}{\partial \gamma} = \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}$$

and

$$\begin{aligned}\frac{\partial \ln G(\gamma, \delta)}{\partial \delta} &= \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)(-v_2)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)(-v_2)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2} \\ &= -\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}.\end{aligned}$$

Then I use (182) to obtain

$$\frac{\partial G(\gamma, \delta)}{\partial \gamma} = G(\gamma, \delta)\frac{\partial}{\partial \gamma}\ln G(\gamma, \delta)$$

and

$$\frac{\partial G(\gamma, \delta)}{\partial \delta} = G(\gamma, \delta)\frac{\partial}{\partial \delta}\ln G(\gamma, \delta).$$

Therefore, the partial derivatives of $G(\gamma, \delta)$ with respect to γ and δ are

$$\frac{\partial G(\gamma, \delta)}{\partial \gamma} = G(\gamma, \delta)\left\{\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right\} \tag{184}$$

$$= 1 - G(\gamma, \delta)\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2} \tag{185}$$

and

$$\frac{\partial G(\gamma, \delta)}{\partial \delta} = G(\gamma, \delta)\left\{-\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right\}. \tag{186}$$

To give the full characterization of the strategic effects of changes in Firm B′s quality on both firms′ prices I substitute (185) and (186) for (107) and use $\frac{\partial\gamma}{\partial r_1} = \frac{1}{q_1 - r_1}[\gamma(q,r) - C_1(r)]$ (equation (172) from the proof of Lemma A) and $\frac{\partial\delta}{\partial r_1} = \frac{1}{q_1 - r_1}\delta(q,r)$ (equation (176) from the proof of Lemma A):

$$
\begin{aligned}
\frac{\partial\widehat{p}_A}{\partial r_1} &= -G(\gamma,\delta) + (q_1 - r_1)\left[\frac{\partial G(\gamma,\delta)}{\partial\gamma}\frac{\partial\gamma}{\partial r_1} + \frac{\partial G(\gamma,\delta)}{\partial\delta}\frac{\partial\delta}{\partial r_1}\right] \\
&= -G(\gamma,\delta) + G(\gamma,\delta)\left\{\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]\right. \\
&\quad [\gamma(q,r) - C_1(r)] \\
&\quad \left. + \left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,v_2\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]\delta(q,r)\right\} \\
&= G(\gamma,\delta)\left\{-1 + \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]\right. \\
&\quad [\gamma(q,r) - C_1(r)] \\
&\quad \left. + \left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]\delta(q,r)\right\},
\end{aligned}
$$

which gives (50).

Similarly, I substitute (185) and (186) for (108) and use $\frac{\partial\gamma}{\partial r_2} = -\frac{C_2(r)}{q_1 - r_1}$ (equation (173) from the proof of Lemma A) and $\frac{\partial\delta}{\partial r_2} = -\frac{1}{q_1 - r_1}$ (equation (177) from the proof of Lemma A):

$$
\begin{aligned}
\frac{\partial\widehat{p}_A}{\partial r_2} &= (r_1 - q_1)\left[\frac{\partial G(\gamma,\delta)}{\partial\gamma}\frac{\partial\gamma}{\partial r_2} + \frac{\partial G(\gamma,\delta)}{\partial\delta}\frac{\partial\delta}{\partial r_2}\right] \\
&= G(\gamma,\delta)\left\{-\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right] C_2(r)\right. \\
&\quad \left. + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right\},
\end{aligned}
$$

which gives (51).

Then, I substitute (185) and (186) for (109) and use $\frac{\partial\gamma}{\partial r_1} = \frac{1}{q_1 - r_1}[\gamma(q,r) - C_1(r)]$ (equation (172) from the proof of Lemma A) and $\frac{\partial\delta}{\partial r_1} = \frac{1}{q_1 - r_1}\delta(q,r)$ (equation (176) from the proof of Lemma A):

$$\frac{\partial \widehat{p}_B}{\partial r_1} = C_1(r) + G(\gamma,\delta)\Bigg\{-1 + \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta) f_2(v_2)\,\mathrm{d}v_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]$$
$$[\gamma(q,r) - C_1(r)]$$
$$+\left[-\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]\delta(q,r)\Bigg\},$$

which gives (52).

Last, I substitute (185) and (186) for (108) and use $\frac{\partial\gamma}{\partial r_2} = -\frac{C_2(r)}{q_1 - r_1}$ (equation (173) from the proof of Lemma A) and $\frac{\partial\delta}{\partial r_2} = -\frac{1}{q_1 - r_1}$ (equation (177) from the proof of Lemma A):

$$\frac{\partial \widehat{p}_B}{\partial r_2} = C_2(r) + G(\gamma,\delta)\Bigg\{-\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right] C_2(r)$$
$$+\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\Bigg\},$$

which gives (53) and completes the proof.

Proof of Proposition 2

Here I characterize the first-order conditions for Firm A and Firm B in addition to finding the equilibrium curve of indifferent consumers. Given the subgame-perfect equilibrium prices $\widehat{p}_A$ and $\widehat{p}_B$ in Stage 2, the equilibrium qualities $\widehat{q}$ and $\widehat{r}$ are mutual best responses:

$$\widehat{q} \equiv (\widehat{q}_1, \widehat{q}_2) = \underset{q=(q_1,q_2)}{\arg\max} \int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\underline{v}_1}^{\widehat{v}_1(v_2;q,\widehat{r})} [v_1\widehat{r}_1 + v_2\widehat{r}_2 - C(\widehat{r})]\,\mathrm{d}F_1\right\}\mathrm{d}F_2$$
$$+\int_{\underline{v}_2}^{\overline{v}_2}\left\{\int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1} [v_1 q_1 + v_2 q_2 - C(q)]\,\mathrm{d}F_1\right\}\mathrm{d}F_2$$

and

$$\widehat{r} \equiv (\widehat{r}_1, \widehat{r}_2) = \underset{r=(r_1,r_2)}{\arg\max} \int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2;\widehat{q},r))\,[\widehat{p}_B(\widehat{q},r) - C(r)]\,\mathrm{d}F_2.$$

Note that

$$\widehat{v}_1(q,r) \equiv \tilde{v}_1(v_2;\widehat{p}_A(q,r),\widehat{p}_B(q,r),q,r) = \frac{\widehat{p}_A(q,r) - \widehat{p}_B(q,r)}{q_1 - r_1} - v_2\frac{q_2 - r_2}{q_1 - r_1}. \quad (187)$$

I begin the proof by finding Firm A's first-order conditions first. I use Leibniz's rule to obtain the first-order derivative of Firm A's payoffs with respect to q_1

$$\begin{aligned}\int_{\underline{v}_2}^{\overline{v}_2}\int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1}&\left[v_1 - C_1\left(q\right)\right]\mathrm{d}F_1\mathrm{d}F_2\\ &-\int_{\underline{v}_2}^{\overline{v}_2}\left[\widehat{v}_1\left(v_2;q,\widehat{r}\right)q_1 + v_2 q_2 - C\left(q\right)\right] f_1\left(\widehat{v}_1\left(v_2;q,\widehat{r}\right)\right)\frac{\partial\widehat{v}_1\left(v_2;q,\widehat{r}\right)}{\partial q_1}\mathrm{d}F_2\\ &+\int_{\underline{v}_2}^{\overline{v}_2}\left[\widehat{v}_1\left(v_2;q,\widehat{r}\right)r_1 + v_2 r_2 - C\left(r\right)\right] f_1\left(\widehat{v}_1\left(v_2;q,\widehat{r}\right)\right)\frac{\partial\widehat{v}_1\left(v_2;q,\widehat{r}\right)}{\partial q_1}\mathrm{d}F_2\end{aligned} \tag{188}$$

I rearrange the first-order derivative in (188) to obtain

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2}\int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1}\left[v_1 - C_1\left(q\right)\right]\mathrm{d}F_1\mathrm{d}F_2\\ &-\int_{\underline{v}_2}^{\overline{v}_2}\left\{\widehat{v}_1\left(v_2;q,\widehat{r}\right)\left(q_1-\widehat{r}_1\right) - \left[C\left(q\right) - C\left(\widehat{r}\right) - v_2(q_2-\widehat{r}_2)\right]\right\}\\ &\qquad f_1\left(\widehat{v}_1\left(v_2;q,\widehat{r}\right)\right)\frac{\partial\widehat{v}_1\left(v_2;q,\widehat{r}\right)}{\partial q_1}\mathrm{d}F_2.\end{aligned} \tag{189}$$

I set (189) equal to zero and evaluate (189) at the equilibrium qualities. By Lemma A the term inside the curly brackets of (189) is zero. This gives the following first-order condition:

$$\int_{\underline{v}_2}^{\overline{v}_2}\int_{\widehat{v}_1(v_2;\widehat{q},\widehat{r})}^{\overline{v}_1}\left[v_1 - C_1\left(\widehat{q}\right)\right]\mathrm{d}F_1\mathrm{d}F_2 = 0. \tag{190}$$

Lastly, I rearrange (190) which gives:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2}\int_{\widehat{v}_1(v_2;\widehat{q},\widehat{r})}^{\overline{v}_1} v_1\mathrm{d}F_1\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right)\right)\mathrm{d}F_2\right]} = C_1\left(\widehat{q}\right). \tag{191}$$

Next, I derive the first-order derivative of Firm A's payoff with respect to q_2. I obtain the first-order derivative of Firm A's payoff by using Leibniz's rule

$$\begin{aligned}\int_{\underline{v}_2}^{\overline{v}_2}\int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1}&\left[v_2 - C_2\left(q\right)\right]\mathrm{d}F_1\mathrm{d}F_2\\ &-\int_{\underline{v}_2}^{\overline{v}_2}\left[\widehat{v}_1\left(v_2;q,\widehat{r}\right)q_1 + v_2 q_2 - C\left(q\right)\right] f_1\left(\widehat{v}_1\left(v_2;q,\widehat{r}\right)\right)\frac{\partial\widehat{v}_1\left(v_2;q,\widehat{r}\right)}{\partial q_2}\mathrm{d}F_2\\ &+\int_{\underline{v}_2}^{\overline{v}_2}\left[\widehat{v}_1\left(v_2;q,\widehat{r}\right)\widehat{r}_1 + v_2 \widehat{r}_2 - C\left(r\right)\right] f_1\left(\widehat{v}_1\left(v_2 q,\widehat{r}\right)\right)\frac{\partial\widehat{v}_1\left(v_2;q,\widehat{r}\right)}{\partial q_2}\mathrm{d}F_2.\end{aligned} \tag{192}$$

I rearrange (192) to obtain

$$\begin{aligned}&\int_{\underline{v}_2}^{\overline{v}_2}\int_{\widehat{v}_1(v_2;q,\widehat{r})}^{\overline{v}_1}\left[v_2 - C_2\left(q\right)\right]\mathrm{d}F_1\mathrm{d}F_2\\ &-\int_{\underline{v}_2}^{\overline{v}_2}\left\{\widehat{v}_1\left(v_2;q,r\right)\left(q_1-r_1\right) - \left[C\left(r\right) - C\left(q\right) - v_2(q_2-r_2)\right]\right\}\\ &\qquad f_1\left(\widehat{v}_1\left(v_2;q,r\right)\right)\frac{\partial\widehat{v}_1\left(v_2;q,r\right)}{\partial q_2}\mathrm{d}F_2.\end{aligned} \tag{193}$$

I set (193) equal to zero and evaluate (193) at the equilibrium prices. Again, by Lemma A the term inside the curly brackets of (189) is zero. This gives the following first-order condition:

$$\int_{\underline{v}_2}^{\overline{v}_2} \int_{\widehat{v}_1(v_2;\widehat{q},\widehat{r})}^{\overline{v}_1} \left[v_2 - C_2\left(\widehat{q}\right)\right] \mathrm{d}F_1 \mathrm{d}F_2 = 0. \tag{194}$$

Lastly, I rearrange (194) which gives:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} \int_{\widehat{v}_1(v_2;\widehat{q},\widehat{r})}^{\overline{v}_1} v_2 \mathrm{d}F_1 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \left[1 - F_1\left(\gamma - v_2\delta\right) \mathrm{d}F_2\right]} = C_2\left(\widehat{q}\right). \tag{195}$$

I move on to the second part of the equilibrium characterization and find Firm B's first-order conditions next. I begin by differentiating Firm B's profits with respect to r_1

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;q,r\right)\right)\left[-C_1\left(r\right)\right] \mathrm{d}F_2 \\
&+\frac{\partial}{\partial r_1}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;q,r\right)\right) \mathrm{d}F_2\right\} \times \left[\widehat{p}_B\left(\widehat{q},r\right) - C\left(r\right)\right] && (196)\\
&+\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{p}\left(\widehat{q},r\right)\right)\right) \mathrm{d}F_2\right\} \frac{\partial \widehat{p}_A}{\partial r_1} \times \left[\widehat{p}_B\left(\widehat{q},r\right) - C\left(r\right)\right], && (197)
\end{aligned}$$

of which the effect via Firm B's own equilibrium price has been ignored due to Lemma 4 (the effect via $\widehat{p}_B$ on firm's own price is second order due to the Envelope theorem). The first two rows in (196) give how private firm's quality affects its cost and demand. The last row in (197) describes the strategic effect of Firm B's quality on Firm A's price.

To obtain the complete partial derivative, I use use $\frac{\partial \widehat{v}_1}{\partial r_1} = \frac{\widehat{v}_1}{q_1 - r_1}$ to obtain[12]

$$\frac{\partial}{\partial r_1}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right) \mathrm{d}F_2\right\} = \frac{1}{q_1 - r_1} \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right) \widehat{v}_1\left(v_2;\widehat{q},r\right) \mathrm{d}F_2 \tag{198}$$

and using $\frac{\partial \widehat{v}_1}{\partial p^A} = \frac{1}{q_1 - r_1}$ gives

$$\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right) \mathrm{d}F_2\right\} = \frac{1}{q_1 - r_1} \int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right) \mathrm{d}F_2. \tag{199}$$

12 Using (44), $\frac{\partial \widehat{v}_1}{\partial r_1} = \frac{\widehat{p}_A - \widehat{p}_B - v_2(q_2 - r_2)}{(q_1 - r_1)^2} = \frac{1}{q_1 - r_1}\left[\frac{\widehat{p}_A - \widehat{p}_B}{q_1 - r_1} - \frac{v_2(q_2 - r_2)}{q_1 - r_1}\right] = \frac{1}{q_1 - r_1}\widehat{v}_1\left(v_2;\widehat{q},r\right)$.

I use (198) and (199) for (196) and (197) to obtain

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\left[-C_1\left(r\right)\right]\mathrm{d}F_2 \\
&\quad+\underbrace{\frac{1}{q_1-r_1}\int_{\underline{v}_2}^{\overline{v}_2}\left[f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]\widehat{v}_1\left(v_2;\widehat{p}\left(\widehat{q},r\right)\right)\mathrm{d}F_2}_{=\frac{\partial}{\partial r_1}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2;\widehat{q},r))\mathrm{d}F_2\right\}}\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right] \\
&\quad+\underbrace{\frac{1}{q_1-r_1}\int_{\underline{v}_2}^{\overline{v}_2}\left[f_1\ \widehat{v}_1\left(v_2;\widehat{p}\left(\widehat{q},r\right)\right]\mathrm{d}F_2}_{=\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\widehat{v}_1(v_2;\widehat{q},r))\mathrm{d}F_2\right\}}\frac{\partial\widehat{p}_A}{\partial r_1}\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right].
\end{aligned} \tag{200}$$

I set the first-order derivative (200) equal to zero and evaluate it at equilibrium qualities. Because in price equilibrium

$$\frac{\widehat{p}_B-C\left(r\right)}{q_1-r_1}=\frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}, \tag{201}$$

I can simplify the first-order condition (200) as follows

$$\begin{aligned}
&\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\left[-C_1\left(r\right)\right]\mathrm{d}F_2 \\
&\quad+\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\widehat{v}_1(v_2;\widehat{q},\widehat{r})\mathrm{d}F_2\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}\right] \\
&\quad+\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2\frac{\partial\widehat{p}_A}{\partial r_1}\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}\right]=0,
\end{aligned} \tag{202}$$

which after re-arranging becomes

$$\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2\left[\frac{\partial\widehat{p}_A}{\partial r_1}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\widehat{v}_1(v_2;\widehat{q},\widehat{r})\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}-C_1\left(r\right)\right]=0. \tag{203}$$

Thus, the first-order condition (203) is

$$\frac{\partial\widehat{p}_A}{\partial r_1}+\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\right)\widehat{v}_1(v_2;\widehat{q},\widehat{r})\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)f_2\left(v_2\right)\mathrm{d}v_2}=C_1\left(r\right), \tag{204}$$

from where I have omitted the factor $F_1(\widehat{v}_1)$.

I derive Firm B's first-order condition with respect to r_2 next. I differentiate

Firm B's profits with respect to r_2

$$\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\left[C_2\left(r\right)\right]\mathrm{d}F_2$$
$$+\frac{\partial}{\partial r_2}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2\right\}\times\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right] \tag{205}$$
$$+\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2\right\}\frac{\partial \widehat{p}_A}{\partial r_2}\times\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right], \tag{206}$$

of which the effect on Firm B's price via its own equilibrium price has been ignored due to Lemma 4 (the effect via $\widehat{p}_B$ on firms' own price is second order). To obtain the complete partial derivative, I use $\frac{\partial \widehat{v}_1}{\partial r_2}=\frac{v_2}{q_1-r_1}$ to obtain

$$\frac{\partial}{\partial r_2}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2\right\}=\frac{1}{q_1-r_1}\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)v_2\mathrm{d}F_2. \tag{207}$$

and use $\frac{\partial \widehat{v}_1}{\partial p^A}=\frac{1}{q_1-r_1}$ to obtain

$$\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2\right\}=\frac{1}{q_1-r_1}\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2. \tag{208}$$

Using (208) and (208) for (205) and (206) gives

$$\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;q,r\right)\right)f_2\left(v_2\right)\left[-C_2\left(r\right)\right]\mathrm{d}F_2$$
$$+\underbrace{\frac{1}{q_1-r_1}\int_{\underline{v}_2}^{\overline{v}_2}\left[f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\right]v_2\mathrm{d}F_2}_{=\frac{\partial}{\partial r_2}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2\right\}}\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right]$$
$$+\underbrace{\frac{1}{q_1-r_1}\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2}_{=\frac{\partial}{\partial \widehat{p}_A}\left\{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1\left(v_2;\widehat{q},r\right)\right)\mathrm{d}F_2\right\}}\frac{\partial \widehat{p}_A}{\partial r_2}\left[\widehat{p}_B\left(\widehat{q},r\right)-C\left(r\right)\right]. \tag{209}$$

I set the first-order derivative (209) equal to zero and evaluate it at equilibrium qualities. Because in price equilibrium

$$\frac{\widehat{p}_B-C\left(r\right)}{q_1-r_1}=\frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}, \tag{210}$$

I can simplify the first-order condition (200) as follows

$$\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)f_2\left(v_2\right)\left[-C_2\left(r\right)\right]\mathrm{d}F_2$$
$$+\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)v_2\mathrm{d}F_2\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2}\right]$$
$$+\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)\mathrm{d}F_2\frac{\partial \widehat{p}_A}{\partial r_2}\left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)f_2\left(v_2\right)\mathrm{d}v_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right)f_2\left(v_2\right)\mathrm{d}v_2}\right]=0,$$

which after re-arranging becomes

$$\int_{\underline{v}_2}^{\overline{v}_2} F_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right) f_2\left(v_2\right) \mathrm{d}v_2 \left[\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right) \mathrm{d}F_2} - C_2\left(r\right)\right] = 0. \tag{211}$$

Thus, the first-order condition with respect to (211) becomes

$$\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right) v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}_1(v_2;\widehat{q},\widehat{r})\right) f_2\left(v_2\right) \mathrm{d}v_2} = C_2\left(r\right), \tag{212}$$

from where I have omitted the factor $F_1(\widehat{v}_1)$. Last, I consider the equilibrium curve of indifferent consumers. Because equilibrium prices $\widehat{p}_A$ and $\widehat{p}_B$ have to follow Lemma 4,

$$v_1\left(v_2;q,r\right) = \frac{C\left(q\right) - C\left(r\right)}{q_1 - r_1} - v_2 \frac{q_2 - r_2}{q_1 - r_1}. \tag{213}$$

This together with equations (191), (195), (204), (212), and (213) give (54)-(57), and (58) in Proposition 2, and completes the proof.

Proof of Result 4

Proof follows closely the steps in Result 2 and is omitted.

Proof of Result 5

Similar to the proof of Result 3, I begin by pointing out that Firm A's equations (equations (56) and (57)) have the same form as the ones in the first best characterizations (equations (6) and (7)) the functions that characterize the equilibrium curve of indifferent consumers and the curve of indifferent consumers in the first best.

Thus, in this proof I compare Firm B's first-order conditions with respect to both quality attributes to the corresponding equations in the first best characterization. As in the proof of Result 3, the equilibrium allocation is efficient if the form of the equations are the same after imposing the assumptions on f_i and the pre-unit production cost of quality. I also use Lemma 6 and Proposition 5 from Laine and Ma (2017) when establishing efficiency.

I begin the proof by simplifying the equilibrium curve of indifferent consumers. Suppose the $\widehat{q}$ that maximizes Firm A's payoffs and the $\widehat{r}$ that maximizes Firm B's payoffs have $\widehat{q}_2 = \widehat{r}_2$. Because $\widehat{q}_2 = \widehat{r}_2$, both firms have the same per-unit production cost of quality $C(q) = c(q_1) + \theta q_1 q_2 + c(q_2)$, and hence $\theta = 0$, $\widehat{v}_1$ given by Lemma 4 becomes:

$$\widehat{v}_1\left(v_2;\widehat{q},\widehat{r}\right) = \frac{c\left(\widehat{q}_1\right) - c\left(\widehat{r}_1\right)}{\widehat{q}_1 - \widehat{r}_1}. \tag{214}$$

Note that $\widehat{v}_1$ given by (214) does not depend on v_2. Also, using the same assumptions on the quality-valuation distributions and the per-unit production cost, the curve of indifferent consumers in the first best will be given by the function:

$$v_1^*\left(v_2; q^{\ell*}, q^{h*}\right) = \frac{c(q_1^{h*}) - c(q_1^{\ell*})}{q_1^{h*} - q_1^{\ell*}}. \tag{215}$$

Next, I simplify the equations for the second quality attributes.

Consider equation (5) from the first best characterization. Using (2), because (214) does not depend on v_2, and the assumption on the per-unit production cost, equation (5) becomes:

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - c_2(q_2^\ell) = 0. \tag{216}$$

Using the same steps as in Result 2, Firm B's first-order derivative in (55) simplifies to become

$$\frac{\int_{\underline{v}_2}^{\overline{v}_2} v_2 \mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} \mathrm{d}F_2} - c_2(r_2) = 0. \tag{217}$$

Equations (216) and (217) have the same form.

I continue with the proof by simplifying the equations for the first quality attributes. I substitute $\widehat{v}_1$ for $G(\gamma, \delta)$, use the result of $\widehat{v}_1$ being independent of v_2, and use the assumption of f_2 being uniform. Thus G becomes $G(\widehat{v}_1) = \frac{F(\widehat{v}_1)}{f(\widehat{v}_1)}$.

Then I simplify the equation for the first best. Using (2), because (214) does not depend on v_2, and the assumption on the per-unit production cost, equation (4) becomes:

$$\frac{\int_{\underline{v}_1}^{\overline{v}_1} \widehat{v}_1 \mathrm{d}F_2}{\int_{\underline{v}_1}^{\overline{v}_1} \mathrm{d}F_2} - c_1(q^\ell). \tag{218}$$

Using Lemma 6 from Laine and Ma (2017) I can rewrite (218) as

$$\widehat{v}_1 - \frac{\int_{\widehat{v}_1}^{\overline{v}_1} f(v)G(v)\mathrm{d}v_1}{F(\widehat{v}_1)G(\widehat{v}_1)} - c_1(q^\ell), \tag{219}$$

in which v_1^* is given by (215).

Then I simplify Firm B's first-order condition with respect to r_1, given by (54). Because f_i is uniform, $\widehat{v}_1$ is given by (214), and because the per-unit production cost of quality is $C(q) = c(q_1) + c(q_2)$, the price response function becomes

$$\begin{aligned}\frac{\partial \widehat{p}_A}{\partial r_1} =& G\left(\widehat{v}\right)\left\{-1 + \left[\frac{\int_{\underline{v}_2}^{\overline{v}_2} f_1\left(\widehat{v}\right)\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2}\left[1 - F_1\left(\widehat{v}\right)\right]\mathrm{d}F_2}\right]\left[\widehat{v} - C_1\left(r\right)\right]\right\} & (220)\\ =& -G\left(\widehat{v}_1\right) + \widehat{v}_1 - c_1\left(r_1\right). & (221)\end{aligned}$$

in which I have used the assumption of $\widehat{v}_1$ being independent of v_2 and f_2 being uniform to move from (220) to (221). After substituting (221) for (54) and using the assumptions of uniform f_1 and $\widehat{v}_1$ being independent of v_2, I simplify the first-order condition as follows

$$\widehat{v} - \frac{G\left(\widehat{v}\right)_1}{2} - c_1\left(r_1\right). \tag{222}$$

Because $G(\widehat{v}_1) = \frac{F(\widehat{v}_1)}{f(\widehat{v}_1)}$. Because f is uniform $G'(\widehat{v}_1) = 1$.

Now, equation (221) has exactly the same form as the LHS of equation (29) in Laine and Ma (2017). Also, equation (219) has exactly the same form as the RHS of equation (20) in Laine and Ma (2017).

Because for a step function f_i, the inverse reverse hazard rate is linear, I can use Proposition 5 from Laine and Ma (2017), which says the equilibrium qualities q_1, r_1 and market shares are first best. Then, because the form of the equations (217) and (216) are the same and all equation conditions have to hold together (equilibrium condition), $\widehat{q}_1$, $\widehat{r}_1$, and $\widehat{r}_2 = \widehat{q}_2$ are the first best. This completes the proof.

3.A.4 Derivations of the equations for the examples

Here I derive the equations for the examples by using the assumptions on consumers' valuations and the per-unit production of quality.

First best

Consumers' valuations on two qualities, v_1 and v_2, are uniformly distributed on the support $[1,2]$. The quality cost function is the following:

$$C(q_1, q_2) = \frac{1}{2}q_1^2 + \theta q_1 q_2 + \frac{1}{2}q_2^2.$$

For notational convenience I re-define the first best curve of indifferent consumers as:

$$v_1^*(v_2; q^\ell, q^h) = \alpha(q^\ell, q^h) - \beta(q^\ell, q^h)v_2,$$

in which the intercept is defined by:

$$\alpha(q^\ell, q^h) = \frac{C(q^h) - C(q^\ell)}{q_1^h - q_1^\ell}$$

and the slope is defined by:

$$\beta(q^\ell, q^h) = \frac{q_2^h - q_2^\ell}{q_1^h - q_1^\ell}.$$

Equation (4) can be simplified as shown by the following steps:

$$\frac{\int_1^2 \left\{ \int_1^{v_1^*\left(v_2;q^\ell,q^h\right)} v_1 \mathrm{d}v_1 \right\} \mathrm{d}v_2}{\int_1^2 F_1\left(v_1^*\left(v_2;q^\ell,q^h\right)\right) \mathrm{d}v_2} = C_1(q^\ell)$$

$$\frac{\int_1^2 \left\{ \int_1^{v_1^*\left(v_2;q^\ell,q^h\right)} v_1 \mathrm{d}v_1 \right\} \mathrm{d}v_2}{\int_1^2 F_1\left(v_1^*\left(v_2;q^\ell,q^h\right)\right) \mathrm{d}v_2} = q_1^\ell + \theta q_2^\ell$$

$$\frac{\frac{1}{2}\int_1^2 \left[(\alpha - v_2\beta)^2 - 1\right] \mathrm{d}v_2}{\int_1^2 (\alpha - v_2\beta - 1)\,\mathrm{d}v_2} = q_1^\ell + \theta q_2^\ell$$

$$\frac{\int_1^2 \left[\alpha^2 - 2\alpha v_2\beta + \beta^2 v_2^2 - 1\right] \mathrm{d}v_2}{2\int_1^2 (\alpha - v_2\beta - 1)\,\mathrm{d}v_2} = q_1^\ell + \theta q_2^\ell$$

$$\frac{\left[2\alpha^2 - 2\tfrac{1}{2}\alpha\beta 2^2 + \tfrac{1}{3}\beta^2 2^3 - 2 - \alpha^2 + 2\tfrac{1}{2}\alpha\beta 1^2 + \tfrac{1}{3}\beta^2 1^3 + 1\right]}{2\left[2\alpha - \tfrac{1}{2}4\beta - 2 - \alpha + \tfrac{1}{2}\beta + 1\right]} = q_1^\ell + \theta q_2^\ell$$

$$\frac{1}{2}\left[\frac{(-1)\left(\alpha^2 - 3\alpha\beta + \dfrac{7}{3}\beta^2 - 1\right)}{(-1)\left(\alpha - \dfrac{3}{2}\beta - 1\right)}\right] = q_1^\ell + \theta q_2^\ell$$

$$\frac{1}{2}\left[\frac{1 - \alpha^2 + 3\alpha\beta - \dfrac{7}{3}\beta^2}{1 - \alpha + \dfrac{3}{2}\beta}\right] = q_1^\ell + \theta q_2^\ell.$$

This is (9).

Similarly, equation (5) can be simplified as shown by the following steps:

$$\frac{\int_1^2 \left\{ \int_1^{v_1^*\left(v_2;q^\ell,q^h\right)} v_2 \mathrm{d}v_1 \right\} \mathrm{d}v_2}{\int_1^2 F_1\left(v_1^*\left(v_2;q^\ell,q^h\right)\right) \mathrm{d}v_2} = C_2(q^\ell)$$

$$\frac{\int_1^2 \left\{ \int_1^{v_1^*\left(v_2;q^\ell,q^h\right)} v_2 \mathrm{d}v_1 \right\} \mathrm{d}v_2}{\int_1^2 F_1\left(v_1^*\left(v_2;q^\ell,q^h\right)\right) \mathrm{d}v_2} = q_2^\ell + \theta q_1^\ell$$

$$\frac{\int_1^2 \left[v_2\left(\alpha - v_2\beta - 1\right)\right] \mathrm{d}v_2}{\int_1^2 \left(\alpha - v_2\beta - 1\right) \mathrm{d}v_2} = q_2^\ell + \theta q_1^\ell$$

$$\frac{\int_1^2 \left[\alpha v_2 - v_2\beta^2 - v_2\right] \mathrm{d}v_2}{\int_1^2 \left(\alpha - v_2\beta - 1\right) \mathrm{d}v_2} = q_2^\ell + \theta q_1^\ell$$

$$\frac{\left[\frac{1}{2}2^2\alpha - \frac{1}{3}\beta 2^3 - \frac{1}{2}2^2 - \frac{1}{2}\alpha 1^2 + \frac{1}{3}\beta 1^3 + \frac{1}{2}1^2\right]}{2\alpha - \frac{1}{2}4\beta - 2 - \alpha + \frac{1}{2}\beta + 1} = q_2^\ell + \theta q_1^\ell$$

$$\frac{(-1)\left(\frac{3}{2}\alpha - \dfrac{7}{3}\beta - \frac{3}{2}\right)}{(-1)\left(\alpha - \dfrac{3}{2}\beta - 1\right)} = q_2^\ell + \theta q_1^\ell$$

$$\frac{1}{2}\left[\frac{3 - 3\alpha + \dfrac{14}{3}\beta}{1 - \alpha + \dfrac{3}{2}\beta}\right] = q_2^\ell + \theta q_1^\ell.$$

This is (10).

Equation (6) can be simplified as shown by the following steps:

$$\frac{\int_1^2 \left\{\int_{v_1^*\left(v_2;q^\ell,q^h\right)}^2 v_1 \mathrm{d}v_1\right\} \mathrm{d}v_2}{\int_1^2 \left[1 - F_1\left(v_1^*\left(v_2;q^\ell,q^h\right)\right)\right] \mathrm{d}v_2} = C_1(q^h)$$

$$\frac{\int_1^2 \left\{\int_{v_1^*\left(v_2;q^\ell,q^h\right)}^2 v_1 \mathrm{d}v_1\right\} \mathrm{d}v_2}{\left[1 - F_1\left(v_1^*\left(v_2;q^\ell,q^h\right)\right)\right.} \mathrm{d}v_2 = q_1^h + \theta q_2^h$$

$$\frac{\frac{1}{2}\int_1^2 \left[4 - (\alpha - v_2\beta)^2\right] \mathrm{d}v_2}{\int_1^2 (2 - \alpha + v_2\beta)\, \mathrm{d}v_2} = q_1^h + \theta q_2^h$$

$$\frac{\int_1^2 \left[4 - \alpha^2 + 2\alpha v_2\beta - \beta^2 v_2^2\right] \mathrm{d}v_2}{2\int_1^2 (2 - \alpha + v_2\beta)\, \mathrm{d}v_2} = q_1^h + \theta q_2^h$$

$$\frac{\left[8 - 2\alpha^2 + 2\frac{1}{2}\alpha\beta 2^2 - \frac{1}{3}\beta^2 2^3 - 4 + \alpha^2 - 2\frac{1}{2}\alpha\beta 1^2 + \frac{1}{3}\beta^2 1^3\right]}{2\left[4 - 2\alpha + \frac{1}{2}4\beta - 2 + \alpha - \frac{1}{2}\beta\right]} = q_1^h + \theta q_2^h$$

$$\frac{1}{2}\left[\frac{(-1)\left(2 - \alpha^2 + 3\alpha\beta - \dfrac{7}{3}\beta^2\right)}{(-1)\left(2 - \alpha + \frac{3}{2}\beta\right)}\right] = q_1^h + \theta q_2^h$$

$$\frac{1}{2}\left[\frac{2 - \alpha^2 + 3\alpha\beta - \dfrac{7}{3}\beta^2}{2 - \alpha + \frac{3}{2}\beta}\right] = q_1^h + \theta q_2^h$$

This is (11).

And equation (7) can be simplified as shown by the following steps:

$$\frac{\int_1^2 \left\{ \int_1^{v_1^*(v_2;q^\ell,q^h)} v_2 \mathrm{d}v_1 \right\} \mathrm{d}v_2}{\int_1^2 \left[1 - F_1\left(v_1^*\left(v_2; q^\ell, q^h\right)\right)\right] \mathrm{d}v_2} = C_2(q^h)$$

$$\frac{\int_1^2 \left\{ \int_1^{v_1^*(v_2;q^\ell,q^h)} v_2 \mathrm{d}v_1 \right\} \mathrm{d}v_2}{\int_1^2 \left[1 - F_1\left(v_1^*\left(v_2; q^\ell, q^h\right)\right)\right] \mathrm{d}v_2} = q_2^h + \theta q_1^h$$

$$\frac{\int_1^2 \left[v_2\left(\alpha - v_2\beta - 1\right)\right] \mathrm{d}v_2}{\int_1^2 \left(2 - \alpha + v_2\beta\right) \mathrm{d}v_2} = q_2^h + \theta q_1^h$$

$$\frac{\int_1^2 \left[\alpha v_2 - v_2\beta^2 - v_2\right] \mathrm{d}v_2}{2\int_1^2 \left(2 - \alpha + v_2\beta\right) \mathrm{d}v_2} = q_2^h + \theta q_1^h$$

$$\frac{\left[\frac{1}{2}2^2\alpha - \frac{1}{3}\beta 2^3 - \frac{1}{2}2^2 - \frac{1}{2}\alpha 1^2 + \frac{1}{3}\beta 1^3 + \frac{1}{2}1^2\right]}{\left[4 - 2\alpha + \frac{1}{2}4\beta - 2 + \alpha - \frac{1}{2}\beta\right]} = q_2^h + \theta q_1^h$$

$$\frac{(-1)\left(\frac{3}{2}\alpha - \dfrac{7}{3}\beta - \frac{3}{2}\right)}{(-1)\left(2 - \alpha + \frac{3}{2}\beta\right)} = q_2^h + \theta q_1^h$$

$$\frac{1}{2}\left[\frac{3 - 3\alpha + \dfrac{14}{3}\beta}{2 - \alpha + \frac{3}{2}\beta}\right] = q_2^h + \theta q_1^h$$

This is (12).

Low first quality attribute in the public firm

Price reaction functions:

I use assumptions of v_1 and v_2 being uniformly distributed on the support $[1,2]$ and the quality cost function $C(q_1,q_2)=\frac{1}{2}q_1^2+\theta q_1 q_2+\frac{1}{2}q_2^2$ to obtain the following

$$\begin{aligned} H(\alpha,\beta) &= \frac{\int_1^2 [1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}{\int_1^2 f_1(\alpha-v_2\beta)\,\mathrm{d}F_2} \\ &= \frac{\int_1^2 [2-\alpha+v_2\beta]\,\mathrm{d}v_2}{\int_1^2 1\mathrm{d}v_2} \\ &= 2-\alpha+\frac{3}{2}\beta. \end{aligned}$$

Also:

$$\begin{aligned} \frac{\int_1^2 f_1(\alpha-v_2\beta)\,v_2\mathrm{d}F_2}{\int_1^2 [1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2} &= \frac{\int_1^2 v_2\mathrm{d}v_2}{\int_1^2 [2-\alpha+v_2\beta]} \\ &= \frac{\frac{3}{2}}{2-\alpha+\frac{3}{2}\beta}. \end{aligned}$$

Note that $f_1'(\alpha-v_2\beta)=0$. Then equation (20) from Lemma 3 becomes:

$$\begin{aligned} \frac{\partial\widehat{p}_A}{\partial r_1} = H(\alpha,\beta)\Bigg\{1+&\left[\frac{\int_1^2 f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_1^2 [1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}+\frac{\int_1^2 f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_1^2 f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right] \\ &[\alpha(q,r)-C_1(r)] \\ -&\left[\frac{\int_1^2 f_1(\alpha-v_2\beta)\,v_2\mathrm{d}F_2}{\int_1^2 [1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}+\frac{\int_1^2 v_2 f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_1^2 f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right]\beta(q,r)\Bigg\} \\ &= \left[2-\alpha+\frac{3}{2}\beta\right]+\alpha-r_1-\theta r_2-\left[2-\alpha+\frac{3}{2}\beta\right]\frac{\frac{3}{2}}{2-\alpha+\frac{3}{2}\beta}\beta \\ &= 2-r_1-\theta r_2. \end{aligned}$$

Similarly, equation (21) from Lemma 3 is:

$$\begin{aligned} \frac{\partial\widehat{p}_A}{\partial r_2} = H(\alpha,\beta)\Bigg\{&\left[-\frac{\int_1^2 f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_1^2 [1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}-\frac{\int_1^2 f_1'(\alpha-v_2\beta)\,\mathrm{d}F_2}{\int_1^2 f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\right]C_2(r) \\ &+\frac{\int_1^2 f_1(\alpha-v_2\beta)\,v_2\mathrm{d}F_2}{\int_1^2 [1-F_1(\alpha-v_2\beta)]\,\mathrm{d}F_2}+\frac{\int_1^2 f_1'(\alpha-v_2\beta)\,v_2\mathrm{d}F_2}{\int_1^2 f_1(\alpha-v_2\beta)\,\mathrm{d}F_2}\Bigg\} \\ &= -r_2-\theta r_1+\left[2-\alpha+\frac{3}{2}\beta\right]\frac{\frac{3}{2}}{2-\alpha+\frac{3}{2}} \\ &= \frac{3}{2}-r_2-\theta r_1. \end{aligned}$$

Quality subgame:

From Proposition 1:

$$\begin{aligned}\frac{\partial \widehat{p}_A}{\partial r_1} + \frac{\int_1^2 f_1(\alpha - v_2\beta)\,\widehat{v}_1 \mathrm{d}F_2}{\int_1^2 f_1(\alpha - v_2\beta)\,\mathrm{d}F_2} &= C_1(r) \\ 2 - r_1 - \theta r_2 + \alpha - \frac{3}{2}\beta &= r_1 + \theta r_2 \\ \frac{1}{2}\left[2 + \alpha - \frac{3}{2}\beta\right] &= r_1 + \theta r_2,\end{aligned}$$

which gives equation (33). From Proposition 1:

$$\begin{aligned}\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_1^2 f_1(\alpha - v_2\beta)\,v_2 \mathrm{d}F_2}{\int_1^2 f_1(\alpha - v_2\beta)\,\mathrm{d}F_2} &= C_2(r) \\ -r_2 - \theta r_1 + \frac{3}{2} + \frac{3}{2} &= r_2 + \theta r_1 \\ \frac{3}{2} &= r_2 + \theta r_1\end{aligned}$$

which gives equation (34).

High first quality attribute in the public firm

Price reaction functions:

I use the assumptions of v_1 and v_2 being uniformly distributed on the support $[1, 2]$ and the quality cost function $C(q_1, q_2) = \frac{1}{2}q_1^2 + \theta q_1 q_2 + \frac{1}{2}q_2^2$ to obtain the following:

$$\begin{aligned}G(\gamma, \delta) &= \frac{\int_1^2 F_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2} \\ &= \frac{\int_1^2 [\gamma - v_2\delta - 1]\,\mathrm{d}v_2}{\int_1^2 1 \mathrm{d}v_2} \\ &= \gamma - \frac{3}{2}\delta - 1.\end{aligned}$$

Also:

$$\begin{aligned}\frac{\int_1^2 f_1(\gamma - v_2\delta)\,v_2 \mathrm{d}F_2}{\int_1^2 F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} &= \frac{\int_1^2 v_2 \mathrm{d}v_2}{\int_1^2 [\gamma - v_2\delta - 1]\,\mathrm{d}v_2} \\ &= \frac{\frac{3}{2}}{\gamma - \frac{3}{2}\delta - 1}.\end{aligned}$$

Because $f_1'(\gamma - v_2\delta) = 0$, equation (50) from Lemma 5 becomes:

$$\begin{aligned}
\frac{\partial \widehat{p}_A}{\partial r_1} &= G(\gamma,\delta)\Bigg\{-1 + \left[\frac{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_1^2 F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_1^2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right] \\
&\quad [\gamma(q,r) - C_1(r)] \\
&\quad + \left[-\frac{\int_1^2 f_1(\gamma - v_2\delta)\,v_2\mathrm{d}F_2}{\int_1^2 F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_1^2 v_2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]\delta(q,r)\Bigg\} \\
&= -\left[\gamma - \frac{3}{2}\delta - 1\right] + \gamma - r_1 - \theta r_2 - \left[\gamma - \frac{3}{2}\delta - 1\right]\frac{\frac{3}{2}}{\gamma - \frac{3}{2}\delta - 1}\delta \\
&= 1 - r_1 - \theta r_2.
\end{aligned}$$

Similarly, equation (51) from Lemma 5 is:

$$\begin{aligned}
\frac{\partial \widehat{p}_A}{\partial r_2} &= G(\gamma,\delta)\Bigg\{-\left[\frac{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_1^2 f_1'(\gamma - v_2\delta)\,\mathrm{d}F_2}{\int_{\underline{v}_2}^{\overline{v}_2} f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\right]C_2(r) \\
&\quad + \frac{\int_1^2 f_1(\gamma - v_2\delta)\,v_2\mathrm{d}F_2}{\int_1^2 F_1(\gamma - v_2\delta)\,\mathrm{d}F_2} - \frac{\int_1^2 f_1'(\gamma - v_2\delta)\,v_2\mathrm{d}F_2}{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2}\Bigg\} \\
&= -r_2 - \theta r_1 + \left[\gamma - \frac{3}{2}\delta - 1\right]\frac{\frac{3}{2}}{\gamma - \frac{3}{2}\delta - 1}\gamma \\
&= \frac{3}{2} - r_2 - \theta r_1.
\end{aligned}$$

Quality subgame:

From Proposition 2

$$\begin{aligned}
\frac{\partial \widehat{p}_A}{\partial r_1} + \frac{\int_1^2 f_1(\gamma - v_2\delta)\,\widehat{v}_1\mathrm{d}F_2}{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2} &= C_1(r) \\
1 - r_1 - \theta r_2 + \gamma - \frac{3}{2}\delta &= r_1 + \theta r_2 \\
\frac{1}{2}\left[1 + \gamma - \frac{3}{2}\delta\right] &= r_1 + \theta r_2,
\end{aligned}$$

which gives equation (54). From Proposition 2:

$$\begin{aligned}
\frac{\partial \widehat{p}_A}{\partial r_2} + \frac{\int_1^2 f_1(\gamma - v_2\delta)\,v_2\mathrm{d}F_2}{\int_1^2 f_1(\gamma - v_2\delta)\,\mathrm{d}F_2} &= C_2(r) \\
\frac{3}{2} - r_2 - \theta r_1 + \frac{3}{2} &= r_2 + \theta r_1 \\
\frac{3}{2} &= r_2 + \theta r_1
\end{aligned}$$

which gives equation (55).

3.A.5 Sufficient second order conditions and equilibrium existence

The characterizations and results in this paper focus on the necessary first-order conditions for social surplus and profit maximization. Because the second order conditions and conditions for the existence of a subgame perfect Nash equilibrium in a two-stage game are analytically difficult to derive in detail, I provide a set of generic conditions in this Appendix.

3.A.5.1 Sufficient second order conditions for the maximization of social surplus (first-best)

In this subsection, I consider the conditions that should be satisfied for the first-best in (3) to have a unique maximum. Recall, that the first best in Section 3.3 consists of a vector of qualities $\left(q^{\ell*}, q^{h*}\right) = \left(q_1^{\ell*}, q_2^{\ell*}, q_1^{h*}, q_2^{h*}\right)$ and an assignment of consumers between the two firms that will be determined by v_1^*. The qualities $\left(q^{\ell*}, q^{h*}\right)$ and v_1^* maximizes social surplus W, that is given by (3). Thus, the necessary condition for the $\left(q^{\ell*}, q^{h*}, v_1^*\right)$ to be a maximum, the first-order partial derivatives with respect to $\left(q_1^\ell, q_2^\ell, q_1^h, q_2^h, \check{v}_1\right)$ are zero. These first-order conditions are given by (66)-(69) and (70) in Appendix 3.A.1.

Then, a sufficient condition for $\left(q_1^{\ell*}, q_2^{\ell*}, q_1^{h*}, q_2^{h*}, v_1^*\right)$ to be the maximum can be studied by the analyzing the symmetric Hessian determinant given by

$$\left|H\right| = \begin{vmatrix} W_{11} & W_{12} & W_{13} & W_{14} & W_{15} \\ W_{21} & W_{22} & W_{23} & W_{24} & W_{25} \\ W_{31} & W_{32} & W_{33} & W_{34} & W_{35} \\ W_{41} & W_{42} & W_{43} & W_{44} & W_{45} \\ W_{51} & W_{52} & W_{53} & W_{54} & W_{55} \end{vmatrix},$$

whose successive principal minors are denoted by

$$\left|H_1\right| = W_{11}, \quad \left|H_2\right| = \begin{vmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{vmatrix}, \left|H_3\right| = \begin{vmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \end{vmatrix}, \left|H_4\right| = \begin{vmatrix} W_{11} & W_{12} & W_{13} & W_{14} \\ W_{21} & W_{22} & W_{23} & W_{24} \\ W_{31} & W_{32} & W_{33} & W_{34} \\ W_{41} & W_{42} & W_{43} & W_{44} \end{vmatrix},$$

and $\left|H_5\right| = \left|H\right|$,
in which W_{ii} denotes the second partial derivative of the social surplus function with respect to each variable $\left(q_1^\ell, q_2^\ell, q_1^h, q_2^h, \check{v}_1\right)$.

The second-order sufficient condition for the maximum is that $\left(q^{\ell*}, q^{h*}, v_1^*\right)$ is a maximum if $\left|H_1\right| < 0$, $\left|H_2\right| > 0$, $\left|H\right| < 0$, $\left|H_4\right| > 0$, and $\left|H_5\right| < 0$ (see Chiang 1984, p. 333). Note that when using this condition, all principal minors must be evaluated at the point in which the first-order derivatives are zero, $\left(q_1^{\ell*}, q_2^{\ell*}, q_1^{h*}, q_2^{h*}, v_1^*\right)$.

As discussed in Appendix 3.A.6 below, the first best is solved from the the system of equations, making sure the constraint $v_1 \in [1,2]$ holds. The first best is the one that yields the highest social surplus. Lastly, I check the solution satisfies the first-order conditions.

3.A.5.2 Existence of a subgame perfect Nash equilibrium in a two-stage game

My analysis of the equilibrium of the game is based on an implicit assumption of the existence of the unique equilibrium in both subgames. Because the general existence is analytically difficult to show, I discuss some of the generic conditions for the existence and uniqueness of the subgame perfect Nash equilibrium of the model.

The game I consider is a two-stage (finite) game of perfect information but with an infinite (continunous) actions in each stage. For these games the existence of the perfect Nash Equilibrium is not as simple as if the game were a game of perfect information with the finite action sets (sometimes called as the finite action games).[13] If the game had a finite actions, a well known result of any finite game of perfect information admitting a perfect Nash equilbrium in pure strategies would hold. The existence of the perfect Nash Equilibrium in pure strategies has been demonstrated by Fudenberg and Levine (1983) in the case of *finite action games* of perfect information with continuous payoffs. As demonstrated by Harris (1985) the existence result of the perfect Nash equilibrium in Fudenberg and Levine (1983) does not directly apply in a class of *infinite action games* of perfect information with continuous payoffs. Moreover, Hellwig and Leininger (1987) show that the subgame-perfect equilibria in pure strategies which are measurable functions exists when the payoff functions are continuous, strategy sets are compact, and constraint correnspondences are continuous. Therefore in case of the game analyzed in this paper, the existence of the subgame-perfect equilibrium could be established by showing that the conditions provided by Hellwig and Leininger (1987) hold.

Last and importantly, the vast majority of my results rely on my numerical examples. As discussed in Appendix 3.A.6, for all numerical examples I show that Hessian is negative definite, the sufficient condition for the equilibrium candidate to be a local maximizer. Then, I check if this equilibrium candidate is a global maximum by using the condition of the global maximum must satisfy the first-order conditions given the other player's strategy. After finding all solutions to a firm's first-order condition at the candidate equilibrium and I show that the equilibrium candidate produces the highest payoffs for both firms, that is the firms do not have an incentive to deviate

3.A.6 Details about the numerical simulations

For Examples 1-3, I take equations (4)-(7) and the equilibrium curve of indifferent consumers and find a numerical solution to the system of equations, making sure

13 The definition of a finite action game is the following: the game is a finite action game if the number of predecessors of a given node is finite, and the number of actions available at any given node is finite (see, for example, Harris 1985).

the constraint $v_1 \in [1, 2]$ holds. The first best is the one that yields the highest social surplus. Lastly, I check the solution satisfies the first-order conditions. This process yields the examples collected in Table 1.

For Examples 4-6, I use the following procedure. I take equations (40)-(43) and the equilibrium curve of indifferent consumers with constraints $q, r > 0$ and $r_1 > q_1$, and then I solve this system of equations numerically, making sure the constraint $v_1 \in [1, 2]$ holds.[14] Then I use the procedure described in Judd et al. (2012). I compute the Hessian H, a $4x4$ matrix which consists of cross-partial derivatives of (40)-(43). Then I check that H is negative definite by checking that $(-1)^k D_k$ for all leading principal minors. For all examples, there is one solution that satisfies these conditions.

Lastly, I check if this solution is the global maximum. The key is that the global maximum must satisfy the first-order conditions given the other player's strategy. This means that I find all solutions to a firm's first-order condition at the candidate equilibrium and then show that the solution produces the highest payoffs for both firms, that is the firms do not have an incentive to deviate. This process yields the examples collected in Table 2.

I use a similar procedure for Examples 7-9 too. I take equations (61)-(64) and the equilibrium curve of indifferent consumers with constraints $q, r > 0$ and $r_1 < q_1$, and then I solve this system of equations numerically, making sure the constraint $v_1 \in [1, 2]$ holds. Then I use the procedure described in Judd et al. (2012). I compute the Hessian H, a $4x4$ matrix which consists of cross-partial derivatives of (40)-(43). Then I check that H is negative definite by checking that $(-1)^k D_k$ for all leading principal minors. Again, for all examples, there is one solution that satisfies these conditions. Last, I check if this solution is the global maximum such as above. This process yields the examples collected in Table 3.

I have also conducted the series of sensitivity analyses for the simulations. All of them point to the following conclusion: with too narrow support ($[0, 1]$ for example) and large θ values the solution yields corner solutions. I have conducted simulations using several different values for θ and the results on differentiation and efficiently remain qualitatively the same.

All numerical computations follow the procedure described above, and the programming conducted by using the Mathematica software are collected in a supplement which is available from the author upon request.

[14] $q, r < 0$ are economically meaningless.

References

Ansari, A., Economides, N., and Steckel, J. 1998. The max-min-min principle of product differentiation. Journal of Regional Science 38 (2), 207–230.

Barigozzi, F. and Ma, C.-t. A. 2018. Product differentiation with multiple qualities. International Journal of Industrial Organization (Forthcoming).

Besley, T. and Malcomson, J. M. 2018. Competition in public service provision: The role of not-for-profit providers. Journal of Public Economics 162, 158–172.

Caplin, A. and Nalebuff, B. 1991. Aggregation and imperfect competition: On the existence of equilibrium. Econometrica 59 (1), 25–59.

Chen, Y. and Riordan, M. H. 2013. Profitability of product bundling. International Economic Review 54 (1), 35–57.

Chiang, A. C. 1984. Fundamental methods of mathematical economics. McGraw-Hill.

Cremer, H., Marchand, M., and Thisse, J.-F. 1991. Mixed oligopoly with differentiated products. International Journal of Industrial Organization 9 (1), 43–53.

De Fraja, G. and Delbono, F. 1990. Game theoretic models of mixed oligopoly. Journal of Economic Surveys 4 (1), 1–17.

Degryse, H. and Irmen, A. 2001. Attribute dependence and the provision of quality. Regional Science and Urban Economics 31 (5), 547–569.

Delbono, F., Denicolo, V., and Scarpa, C. 1991. Quality choice in a vertically differentiated mixed duopoly. Working Papers 108. Dipartimento Scienze Economiche, Universita' di Bologna.

Fudenberg, D. and Levine, D. 1983. Subgame-perfect equilibria of finite- and infinite-horizon games. Journal of Economic Theory 31 (2), 251–268.

Gabszewicz, J. J. and Thisse, J.-F. 1979. Price competition, quality and income disparities. Journal of Economic Theory 20 (3), 340–359.

Garella, P. G. and Lambertini, L. 2014. Bidimensional vertical differentiation. International Journal of Industrial Organization 32, 1–10.

Grilo, I. 1994. Mixed Duopoly under Vertical Differentiation. Annales d'Economie et de Statistique (33), 91–112.

Harris, C. 1985. Existence and characterization of perfect equilibrium in games of perfect information. Econometrica 53 (3), 613–628.

Hellwig, M. and Leininger, W. 1987. On the existence of subgame-perfect equilibrium in infinite-action games of perfect information. Journal of Economic Theory 43 (1), 55–75.

Irmen, A. and Thisse, J.-F. 1998. Competition in multi-characteristics spaces: Hotelling was almost right. Journal of Economic Theory 78 (1), 76–102.

Judd, K. L., Renner, P., and Schmedders, K. 2012. Finding all pure-strategy equilibria in games with continuous strategies. Quantitative Economics 3 (2), 289–331.

Laine, L. and Ma, C.-t. A. 2017. Quality and competition between public and private firms. Journal of Economic Behavior & Organization 140, 336–353.

Lauga, D. and Ofek, E. 2011. Product positioning in a two-dimensional vertical differentiation model: The role of quality costs. Marketing Schience 30 (5), 903–923.

Moorthy, K. S. 1988. Product and price competition in a duopoly. Marketing Science 7 (2), 141–168.

Rosenman, R. and Munoz-Garcia, F. 2017. Nonprofit product differentiation. Economics Bulletin 37 (3), 1415–1420.

Shaked, A. and Sutton, J. 1982. Relaxing price competition through product differentiation. The Review of Economic Studies 49 (1), 3–13.

Shaked, A. and Sutton, J. 1983. Natural oligopolies. Econometrica 51 (5), 1469–1483.

Spence, M. 1975. Monopoly, Quality, and Regulation. Bell Journal of Economics 6 (2), 417–429.

Stenbacka, R. and Tombak, M. 2018. Optimal reimbursement policy in health care: Competition, ownership structure and quality provision. The B.E. Journal of Economic Analysis & Policy 18 (1), 1–19.

Stole, L. 2007. Price Discrimination and Competition. In ed. by Armstrong, M. and Porter, R. 1st ed. Vol. 3. Elsevier. Chap. 34, 2221–2299.

Tabuchi, T. 1994. Two-stage two-dimensional spatial competition between two firms. Regional Science and Urban Economics 24 (2), 207–227.

Tirole, J. 1988. The Theory of Industrial Organization. MIT Press.

Vandenbosch, M. B. and Weinberg, C. B. 1995. Product and price competition in a two-dimensional vertical differentiation model. Marketing Science 14 (2), 224–249.

4 HEALTH CARE PROVIDER ENTRY AND QUALITY COMPETITION UNDER REGULATED PAYMENTS AND INACCURATE QUALITY PERCEPTIONS

Abstract[*]
I study the regulation of health care payment schemes for health care providers' entry decisions and quality choices when some patients have inaccurate quality perceptions that affect their decision utilities. The regulator uses prospective payment and cost reimbursement as its payment instruments to reimburse health care services produced by locally competitive providers. A pure prospective payment implements the first best quality if the providers' entry is contractible (by using for example entry licenses). If the providers' entry is not contractible, a mixed payment scheme is needed to establish efficiency. Inaccurate quality perceptions can affect the strength of competition and regulatory policies.

Keywords: Quality, regulation, entry, quality misperception, mixed payment schemes

[*] I thank Anirban Basu, Ari Hyytinen, Fahad Khalil, and Ching-to Albert Ma for their comments. I undertook some of this research when I visited University of Washington; I am grateful for their generosity. I wish to thank Yrjö Jahnsson Foundation for financial support.

4.1 Introduction

Health care reforms aim at promoting efficiency of health care markets and improving health care quality by increasing provider competition (Gaynor et al. 2015; Pauly et al. 2011). As a part of the many recent and ongoing reforms in Europe and the United States, potential entry barriers are removed to enhance provider entry, and patients are allowed to choose their preferred medical provider more freely. However, the supply and demand responses to these policies can be complicated. Removing entry regulations can lead to under- and over-entry (Amir et al. 2014; Mankiw and Whinston 1986). Regarding the effects of competition on health care quality, even in the competitive private markets with goods priced at marginal costs, quality is inefficient (Spence 1975).

When health care is reimbursed by an insurer, patients' choice is driven by non-price characteristics such as quality. Thus, a demand response requires that patients know what health care quality is. Even if patients know the quality information perfectly, providers may react to the increased demand differently: for example, there can be over-provision to low-severity patients or under-provision to high-severity patients (skimping) or even avoidance of high-severity patients (dumping) (Ellis 1998). Information on provider quality can be difficult to obtain and process.[1] The choice procedure can be difficult too, as illustrated in the markets for other complex products, such as financial investments and health insurance plans (Abaluck and Adams 2017; Abaluck and Gruber 2011; Handel and Kolstad 2015).

This paper studies the effects of provider payment scheme regulation for health care provider entry and health care quality, and further it looks at how this is affected by the competition environment and patients' misperceptions of health care quality. I analyze a model in which locally competitive health care providers, such as hospitals, groups of physicians, or individual physicians, make strategic choices on entry and health care quality. The model builds on Salop (1979) with a few extensions. First, instead of a price, the providers choose the quality of the medical good or service that they offer. Second, the regulator covers all patients' medical expenses, and the only cost patients face is a transportation cost to the provider's location. Third, the regulator makes a choice on how the payment scheme is regulated. Fourth, some patients may overreact or underreact to the perceived health care quality, which I call quality misperception or responsiveness to the perceived quality. To my knowledge, this model is the first that combines provider entry and quality choices, quality misperception, and payment regulation, all typical characteristics of health care reforms.

I study two separate regulatory frameworks. In the first framework the regulator uses entry regulation to control for the number of providers in the market and uses prospective payment as its payment instrument. This regulatory

1 Because health care is a complex product, patients may have incorrect information about providers' quality and services (Arrow 1963). Empirically, the correlation between patient experience and technical quality of care has been found to be mixed (Doyle et al. 2013).

framework is interesting because entry regulation is commonly used in different markets (pharmacy market and medical licensing in health care are examples of these). In the second regulatory framework, the number of providers is not contractible, and the regulator uses a mixed payment instrument, a combination of prospective payment and cost reimbursement, as its payment scheme. When the regulator uses prospective payment, it pays a fixed price for delivering a medical care service irrespective of resources used.[2] When the regulator uses cost reimbursement, the health care provider receives some revenue corresponding to resources that were used to produce medical care. Because the regulator covers all medical expenses for the patient, this model can be applied to health care markets characterized by price regulation when it forces providers to charge the same regulated price for a given health care service. This applies to many markets, such as many European health care systems, Medicare, and Medicaid programs in the United States.

The timing of my model is as follows. In the first stage, the regulator chooses a payment regulation scheme. In the second stage, the providers observe the payment scheme and make an entry choice. In the third stage, the providers choose their qualities. Then patients observe these qualities and choose their provider. Following the terminology originating from Kahneman et al. (1997), there is a share of patients that misperceive health care quality.[3] Unlike using a more general formulation for patient misperception, such as used by Brekke et al. (2012) and Mak (2017), I model quality misperception by allowing some patients to be more or less responsive to quality than the others: patients may overreact or underreact to quality. This simple misperception formulation allows me to study how equilibrium entry and qualities are affected by patient misperception in a tractable way.

I find that entry regulation combined with pure prospective payment allows the regulator to implement first best quality. When the number of providers is regulated, the regulator fixes the number of entering providers to the first best level and may be able to choose the amount of the prospective payment such that it implements the first best quality. To my knowledge my article is the first documenting this. The result is similar to the classic finding on provider payment scheme regulation by Ma (1994): the contractible degree of horizontal differentiation combined with a pure prospective payment scheme leads to the first best outcome on quality. Because Ma (1994) focuses on imperfectly competitive markets and does not consider entry, my result complements the existing literature with the finding that in competitive markets the first best outcome may be obtained by a pure provider payment combined with entry licensing.

In this case, the regulator may be able to the level of the prospective payment and the cost reimbursement such that the equilibrium number of providers and quality is efficient (in other words when the model parameters are such that

---

2 Diagnosis-related groups (DRG-payments) for inpatient hospital services and capitation payments are examples of prospective payments.

3 This is different to the concept where patients are misinformed about the probabilities of product failure, which means that the consumer is misinformed about the product.

there is neither full cost sharing nor no cost sharing). This result is similar to Bardey et al. (2012), who find that a mixed reimbursement scheme is welfare improving in imperfectly competitive markets with horizontally and vertically differentiated products. My article complements the findings in Bardey et al. (2012) by showing that in an equilibrium a mixed payment scheme could be used to yield first best entry and quality outcomes in locally competitive markets too.

If the regulator is constrained to choosing its payment scheme in a regime where the cost reimbursement constraint binds, that is when the model parameters are such that there is either full cost sharing or no cost sharing, the equilibrium outcome is inefficient. The regime in which the regulator is constrained to choosing its payment scheme depends on the exogenous parameters of the model: the competition environment (transportation costs and entry costs) and the (variable) costs of producing quality.

Fiercer competition or increased cost sharing because of higher provider payments increases equilibrium quality.[4] The intuition is the following. A greater number of providers weakens the cost containment effect and leads to increased quality.[5] Because a decrease in transportation cost decreases the degree of differentiation, patients react more to changes in quality, which gives an individual provider an incentive to increase its quality. Increasing cost sharing also leads to increased equilibrium quality: it increases the share of the reimbursed costs of quality production and thus increases the per-unit profit margin. However, the channel through which the payment scheme regulation affects cost sharing depends on which payment instrument is used: an increase in prospective payment increases the per-unit sales of a provider, but an increase in cost reimbursement reduces its per-unit costs.

The effects of the competition environment and the payment scheme regulation on provider entry are as follows. Unsurprisingly, the lower the entry cost is, the larger number of providers enter. However, fiercer competition through decreased transportation costs reduces provider entry. When the transportation cost decreases, the degree of differentiation decreases. This leads to fiercer post-entry competition and reduced provider entry. I also find that the effects of both payment instruments are, in general, ambiguous. However, the effects of both payment instruments are positive if the indirect cost effect, caused by a higher equilibrium quality, is smaller than the direct effect on profitability.

I find that the equilibrium quality is increasing in the responsiveness to the perceived quality (quality misperception), and increased responsiveness to quality decreases provider entry. A way to interpret the responsiveness to the quality is describe it as a determinant of the quality elasticity of the demand of those patients who perceive quality inaccurately. Thus, ceteris paribus, the elasticity

4 Competition becomes fiercer when the number of providers increases or the degree of differentiation (transportation cost) decreases.

5 The cost containment effect here has similar features to the cost containment effect in Bardey et al. (2012). In their model, the cost containment emerges from location changes of providers and market shares. In my model, the cost containment effect arises from increased entry of providers, which are allocated exogenously and located symmetrically in the market.

is higher when quality mispeception parameter is higher, implying even a small change in quality can cause a large shift in demand. Also, increasing the responsiveness to quality increases the quality elasticity of demand of those patients who perceive quality imperfectly which decreases provider entry. In other words, patients are very responsive to quality and demand becomes more elastic. More elastic demand leads to fiercer competition, lower profit margins, and reduced entry.

This article contributes to the large literature on the provider payment scheme regulation. In particular, the article complements the literature originating from Ma (1994) studying the effects of provider payment scheme regulation in markets with product differentiation.[6] My article is closely related to the article by Bardey et al. (2012). They analyze the effects of payment scheme regulation (prospective payments and cost reimbursement) on quality efficiency using a duopoly model where the health care service quality provided is horizontally and vertically differentiated.[7] I instead focus on entry decisions and patient misperception in competitive markets.

As far as I am aware of, the only paper focusing on provider payment scheme regulation and entry is by Waterson (1993). In addition to an empirical analysis on the retail pharmacy market in Melbourne, he provides a model that builds on Salop (1979) and studies pharmaceutical entry to local markets when pharmacies compete on prices. He considers regulation of drug prices and assumes a particular form for the total cost function. My focus instead is on competition for local health care provision: providers compete on the quality of provision as opposed to prices; the quality production function is more general; some patients misperceive quality; and the regulator uses prospective payments and cost-reimbursement of quality.

Last, I contribute to the literature on quality misperception and the strategic behavior of firms (see for example recent papers by Brekke et al. 2012; Mak 2017). I complement this literature by focusing on the effects of payment scheme regulation and patient misperception on health care provider entry and quality choices in locally competitive markets.

The rest of the article is structured as follows. Section 4.2 describes the model. Section 4.3 characterizes the symmetric subgame-equilibrium qualities and the equilibrium number of providers. Section 4.4 characterizes the first best and studies the two regulatory frameworks. Section 4.5 discusses the policy implications. The last section concludes. Appendix 4.A. collects the proofs, derivations, and provides an example with a more specific form of the production costs.

---

6 Newhouse (1996), Ellis et al. (2017), and Ma and Mak (2018) provide excellent summaries on the literature of provider payment incentives and regulation.

7 Bardey et al. (2012) pin down exactly the regime quality is efficient; that is if the quality equilibrium induced by using prospective payment only is characterized by too low quality (vertical differentiation) and overspecialization (horizontal differentiation).

4.2 The Model

4.2.1 Patients

There is a set of patients uniformly distributed on a unit-circumference circle. A patient x demands one unit of a health care service or a good from one of the local health care providers i. The benefit of the service consists of the sum of a common benefit of the health care service $\bar{s}$ and a provider-specific health care quality q_i. I also assume $\bar{s}$ is large enough so that the market is covered.

All medical costs related to the service are covered by a third party payer which I call the regulator. The only cost a patient faces is the linear transportation cost τ. τ measures the distance of how far the patient is from the ideal location. In the literature τ is often interpreted to measure competition intensity in the market (see for example Tay (2003), who studies patients' choice of hospitals).

Some patients may overreact or underreact to the perceived health care quality.[8] I call this *quality misperception*. Patients who perceive quality perfectly are called *patients with accurate perceptions*. Their share of all patients is $\lambda > 0$. Patients who perceive quality imperfectly are called *patients with inaccurate perceptions*. Their share of all patients is $1 - \lambda$. At times, I refer to a higher value of λ by saying that there is more information available to patients about the provider quality. Parameter $a > 0$ measures how *responsive* patients are to quality. If $a > 1$, patients *overreact* to quality. If $0 < a < 1$, patients *underreact* to quality. If a is high, patients who perceive quality imperfectly are very responsive to quality. The average degree of misperception in the market is $\kappa(\lambda, a) = \lambda + (1 - \lambda) a$. If κ is very close (or equal) to one, the average degree of misperception is low. The further κ is from one, the greater the average degree of misperception is.[9]

For this behavioral bias I follow terminology from Kahneman et al. (1997) and let the experienced utility of a service for some patients to differ from their decision utility of this service. Thus, formally $\bar{s} + q$ represents the true, experienced benefit from a service. Social welfare is based on the experienced benefit. Patients make choices based on their decision utilities. The decision utility is the sum of the decision benefit of a good and the costs. For patients with accurate perceptions, the experienced benefit of a service is the same as their decision benefit $\bar{s} + q$. For the patients with inaccurate perceptions the experienced benefit is different from their decision benefit. For them, this decision benefit is $\bar{s} + aq$.

---

8 Patients do not misperceive τ. This assumption is natural for example when τ is interpreted as the physical location of the provider.

9 Strictly speaking λ and a can be interpreted as capturing the same phenomenon, and thus λ would be zero without a loss of generality. However, introducing a and λ separately allows me to distinguish further the effects of responsiveness to quality a from the effects of changes in the share of patients with accurate quality perception λ.

4.2.2 Providers and the regulator

There is a large number of potential profit-maximizing health care providers.[10] A profit-maximizing provider decides whether to enter the market or not. Entering the market requires a fixed cost $K > 0$, and this cost is sunk after the provider has entered. If a provider decides not to enter, its payoffs are 0. After entering the market, n symmetrically located providers compete in quality.

The production of quality consists of a fixed cost of producing quality denoted by $C(q)$ and a variable cost of producing quality c. The difference between the two cost components is that the former is fixed in the sense that it does not vary with the amount of services provided, whereas the latter has to be paid for every unit that a provider sells. To clearly separate between K and $C(q)$, I call the former fixed cost of entry and the latter fixed cost of producing quality. For $C(q)$, I assume that $\frac{\partial C(q)}{\partial q} > 0$ and $\frac{\partial^2 C(q)}{\partial^2 q} > 0$. Also, $C(0) = 0$, $lim_{q \to 0} C'(q) = 0$, and $lim_{q \to \infty} C'(q) = +\infty$.

For each unit of a health care service supplied, a provider receives a non-negative prospective payment p and a cost reimbursement $\alpha \in [0, 1]$. The cost reimbursement α is based on the quality-related variable costs c that the provider reports to the regulator.[11] Then the total payment per patient is $p - (1 - \alpha)c$, and the profit of a provider i is $\pi_i = [p - (1 - \alpha)cq_i] \, D_i \left(q_i, q_{-i}\right) - C\left(q_i\right)$, with demand $D_i \left(q_i, q_{-i}\right)$, and q_{-i} is the quality of the closest competitors $i + 1$ and $i - 1$. I derive the demand function, $D_i(q_i, q_{-i})$, explicitly below. I also assume that the model parameters are such that it is always profitable for at least one provider to enter the market. This means that the post-entry profits are weakly positive in a monopoly by assumption.

The social welfare maximizing regulator chooses a payment regulation scheme $R = (p, \alpha)$. Social welfare is based on the experienced (true) benefit of the service provided to the patient. The regulator does not reimburse the fixed quality production costs of quality or the entry cost. Last, to ensure that the equilibrium quality is strictly positive for all n, $n > 1$, the regulator chooses p always so that $p > \frac{c\tau}{\kappa(\lambda, a)}$.

4.2.3 Extensive form

I study the symmetric subgame-perfect equilibria of the following game:

Stage 1: The regulator chooses and commits to the payment regulation scheme which maximizes the utilitarian social welfare.

[10] I focus on pure private markets and do not consider payment scheme regulation of public and private providers. Analyzing public-private competition would introduce non-symmetric objectives for the firms and thus require different type of modeling approach than Salop.

[11] It is natural to assume that the regulator is not able to reimburse the fixed costs of producing quality. One argument for this is that variable costs are typically easier to measure and verify.

Stage 2: Providers observe the payment scheme and choose whether to enter or not. If a provider decides to enter, it incurs a fixed entry cost. This stage defines the equilibrium number of providers.

Stage 3: Providers observe the number of providers that have entered the market. Providers choose the level of quality they provide. This defines the equilibrium vector of qualities.

Stage 4: Qualities are revealed to the patients. A patient picks a provider based on their experienced benefit of the service.

The outcome of the game consists of a vector of qualities $Q = (q_1, ..., q_n)$, the number of providers n, the payment scheme $R = (p, \alpha)$, and the allocation of patients across providers. I describe patient behavior in the next subsection before moving on to the equilibrium characterization.

4.2.4 Patient behavior

The decision problem of the patient x is to maximize the decision utility obtained from the health care service. If a patient x has *accurate* quality perceptions, the purchasing decision between providers i and $i+1$ $\max_{k=i,i+1} \{\bar{s} + q_k - \tau |l_k - x|\}$. If the patient x has *inaccurate* quality perceptions, the purchasing decision between providers i and $i+1$ is defined formally by $\max_{k=i,i+1} \{\bar{s} + aq_k - \tau |l_k - x|\}$.

Given the payment scheme $R = (p, \alpha)$, the number of providers n, and qualities $Q = (q_1, ..., q_n)$, the patient with *inaccurate* quality perceptions who is indifferent between providers i and $i+1$ is $\widetilde{x}^{inac}_{i,i+1}$ as defined by $\bar{s} + aq_i - \tau\big(\widetilde{x}^{inac}_{i,i+1} - \frac{i}{n}\big) = \bar{s} + aq_{i+1} - \tau\big(\frac{i+1}{n} - \widetilde{x}^{inac}_{i,i+1}\big)$. Solving $\widetilde{x}^{inac}_{i,i+1}$ gives

$$\widetilde{x}^{inac}_{i,i+1} = \frac{a\,(q_i - q_{i+1})}{2\tau} + \frac{2i+1}{2n}.$$

With the same analogy, the patient with *inaccurate* quality perceptions who is indifferent between providers i and $i-1$ is $\widetilde{x}^{inac}_{i-1,i}$ is defined by $\bar{s} + aq_i - \tau\big(\widetilde{x}^{inac}_{i-1,i} - \frac{i}{n}\big) = \bar{s} + aq_{i+1} - \tau\big(\frac{i+1}{n} - \widetilde{x}^{inac}_{i-1,i}\big)$. Solving $\widetilde{x}^{inac}_{i-1,i}$ gives

$$\widetilde{x}^{inac}_{i-1,i} = \frac{a\,(q_{i-1} - q_i)}{2\tau} + \frac{2i-1}{2n}.$$

Last, the patient with *accurate* quality perceptions who is indifferent between providers i and $i+1$ is $\widetilde{x}^{ac}_{i+1,i}$ is defined by $\bar{s} + q_i - \tau\big(\widetilde{x}^{ac}_{i,i+1} - \frac{i}{n}\big) = \bar{s} + q_{i+1} - \tau\big(\frac{i+1}{n} - \widetilde{x}^{ac}_{i,i+1}\big)$. Solving $\widetilde{x}^{ac}_{i+1,i}$ gives

$$\widetilde{x}^{ac}_{i,i+1} = \frac{q_i - q_{i+1}}{2\tau} + \frac{2i+1}{2n}.$$

A patient with *accurate* quality perceptions who is indifferent between providers i and $i-1$ is $\widetilde{x}^{ac}_{i-1,i}$ is determined by $\bar{s} + q_i - \tau\big(\widetilde{x}^{ac}_{i-1,i} - \frac{i}{n}\big) = \bar{s} + q_{i+1} - \tau\big(\frac{i+1}{n} - \widetilde{x}^{ac}_{i-1,i}\big)$. Solving $\widetilde{x}^{ac}_{i-1,i}$ gives

$$\widetilde{x}^{ac}_{i-1,i} = \frac{q_{i-1} - q_i}{2\tau} + \frac{2i-1}{2n}.$$

The demand for provider i from patients with inaccurate quality perceptions is

$$D_i^{inac}\left(q_i, q_{-i}\right) = \frac{a\left(q_i - q_{-i}\right)}{\tau} + \frac{1}{n}, \tag{1}$$

in which q_{-i} refers to the qualities of the closest competitors $i+1$ and $i-1$. Similarly, the demand for provider i from patients with accurate quality perceptions is

$$D_i^{ac} = D_i^{ac}\left(q_i, q_{-i}\right) = \frac{q_i - q_{-i}}{\tau} + \frac{1}{n}, \tag{2}$$

in which q_{-i} refers to the qualities of the closest competitors $i+1$ and $i-1$. The total demand faced by provider i is

$$D_i = D_i\left(q_i, q_{-i}\right) = \lambda D_i^{ac}(q_i, q_{-i}) + (1-\lambda)\, D_i^{inac}\left(q_i, q_{-i}\right). \tag{3}$$

Thus, substituting (1) and (2) for (3) gives:

$$D_i = D_i\left(q_i, q_{-i}\right) = \lambda\left\{\frac{q_i - q_{-i}}{\tau}\right\} + (1-\lambda)\left\{\frac{a\left(q_i - q_{-i}\right)}{\tau}\right\} + \frac{1}{n}. \tag{4}$$

4.3 Equilibrium

4.3.1 Quality choices

Consider the subgame in Stage 3, which is defined by a payment scheme $R = (p, \alpha)$ and a number of health care providers n. The profits of a provider i are

$$\pi_i = \pi_i\left(q_i, q_{-i}\right) = \left[p - (1-\alpha)cq_i\right] D_i\left(q_i, q_{-i}\right) - C\left(q_i\right). \tag{5}$$

A health care provider i chooses its q_i to maximize (5), given the total demand (4) and competitors' quality q_{-i}. Equilibrium qualities $Q = \left(q_1^*, ..., q_i^*, ..., q_n^*\right)$ are best responses against other providers' qualities.

Differentiating (5) with respect to q_i and setting this first-order derivative to zero gives the following first-order condition:

$$\left[p - (1-\alpha)cq_i\right]\frac{\partial D_i\left(q_i, q_{-i}\right)}{\partial q_i} - (1-\alpha)cD_i\left(q_i, q_{-i}\right) - C'\left(q_i\right) = 0 \tag{6}$$

The first term on the left-hand side of (6) shows how a higher quality increases provider i's revenue. The second term in the left-hand side of (6) captures the increased variable costs of producing higher quality. The last term is the increase in the fixed, volume-independent, costs of producing higher quality. These terms reveal a trade-off between a higher market share and the associated quality costs: an increase in quality increases the demand and increases the provider's market share. This increases profits. However, increasing quality also increases associated quality costs. This decreases profits.

In a symmetric equilibrium, providers are located equidistantly. After I impose symmetry $q = q_i = q_{-i}$, the demand in (4) for provider i becomes $1/n$. Thus, the first-order condition (6) becomes

$$G(q;z) \equiv [p - (1-\alpha)cq] \frac{\kappa(\lambda, a)}{\tau} - \frac{(1-\alpha)c}{n} - C'(q) = 0, \tag{7}$$

in which $z = (p, \alpha, c, n, \tau, a, \lambda)$, and $\kappa(\lambda, a) = \lambda + (1-\lambda)a$. The first-order condition in (7) implicitly defines q as a function of $z = (p, \alpha, c, n, \tau, a, \lambda)$. The following proposition states the result for equilibrium qualities.[12]

Proposition 1 *In subgame (n, R), the symmetric subgame equilibrium quality q^* is determined implicitly by equation (7).*

Proposition 1 shows that the equilibrium quality depends on the competition parameters (n, τ), the regulation parameters, $R = (p, \alpha)$, as well as the misperception parameters (a, λ). The proposition also implies that if the symmetric equilibrium characterized by $G(q^*; z) = 0$ exists, it must be that the per-unit profits are positive, in other words, $p - (1-\alpha)cq^* > 0$ in equilibrium. Using this feature and applying the implicit function theorem around q^* gives the comparative statics of equilibrium quality with respect to the competition, regulation, and misperception parameters. I begin with the competition parameters and obtain the following:

Corollary 1 *Equilibrium quality q^* is increasing in n and decreasing in τ.*

Corollary 1 implies that fiercer competition leads to a higher equilibrium quality. Competition becomes fiercer when either the number of providers increases or the degree of differentiation decreases. I explain this result as follows. A greater number of providers implies a lower market share for each provider. The resulting lower market share decreases the variable costs of providing quality and hence induces each provider to increase its quality. A greater number of providers thus weakens the cost containment effect. The cost containment effect here is similar - but not identical - to the cost containment effect in Bardey et al. (2012). In their model of horizontal and vertical differentiation in a private duopoly, the cost containment emerges from a location choice of providers and market shares. Here, the cost containment effect arises from increased entry of providers, which are allocated exogenously and located symmetrically in the market.

The degree of differentiation decreases when the transportation cost decreases. Decreased differentiation leads to a higher quality because then, ceteris paribus, patients react more to changes in quality. Thus a decrease in the degree of differentiation gives an individual provider an incentive to increase its quality.

The following characterizes how equilibrium quality is affected by payment scheme regulation.

---

[12] Note that when $p > \frac{c\tau}{\kappa(\lambda,a)}$ (as I have assumed), $G(0,z) = \frac{p\kappa}{\tau} - \frac{(1-\alpha)c}{n} > 0$ for all $n \geq 1$ and all $0 \leq \alpha \leq 1$. This implies that the providers choose a strictly positive quality in the symmetric equilibrium.

Corollary 2 *Equilibrium quality q^* increases in p and α.*

By increasing either of the two payment instruments, the regulator reimburses a greater share of the costs of quality production. An increase in the prospective payment increases the per-unit sales of a provider, whereas an increase in the cost reimbursement reduces its per-unit costs. Both of these changes increase the per-unit profit margin of attracting more patients and thus give an incentive to increase quality.

The following presents how patient misperception affects equilibrium quality.

Corollary 3 *Equilibrium quality q^* increases in a. If patients overreact to quality, equilibrium quality q^* decreases in λ. The effect is the opposite if patients underreact to quality.*

A way to interpret the responsiveness to quality, a, is to think it as a determinant of the quality elasticity of the demand of those patients who perceive quality inaccurately. In the symmetric equilibrium, the quality elasticity of the demand of *all* patients is $\frac{nq^*}{\tau(\lambda+(1-\lambda)a)}$, which shows that, ceteris paribus, the elasticity is higher when parameter a is higher. Thus, when the responsiveness to quality is high, even a small change in quality can cause a large shift in demand.

The effect of increasing the share of patients with accurate quality perception depends on whether patients underreact or overreact to quality. If patients overreact to perceived quality, an increase in the share of patients with accurate quality perception decreases the equilibrium quality. The intuition is the following. When patients with inaccurate perception overreact to quality changes ($a > 1$) and when the share of patients with accurate quality perception increases, the demand becomes less responsive to changes in quality. This reduction leads to a decrease in the equilibrium quality, because the incentives of providers to compete in qualities are lessened. The effect of an increase in the share of patients with accurate quality perception is exactly the opposite when patients with inaccurate perception underreact to quality changes ($a < 1$).

I turn now to study if patient misperception alleviates or strengthens the effects of competition and payment scheme regulation. These are given by the cross-partial derivatives of the competition and regulation parameters with respect to a and λ and are summarized in the following corollary.

Corollary 4 *The following results are obtained provided that $C''' = 0$.*
i) Competition parameters: An increase in a weakens the effect entry, n, on q^*. Moreover, if some patients overreact to quality ($a > 1$), an increase in λ strengthens the effect of n. If some patients underreact to quality ($a < 1$), the effect of λ is the opposite. The effects of a and λ on the effect of τ are ambiguous.
ii) Regulatory parameters: An increase in a strengthens the effect of prospective payment on q^*. Moreover, if some patients overreact to quality ($a > 1$), increase in λ weakens the effect of p on q^*. If some patients underreact to quality ($a < 1$), the effect of λ is the opposite. The effects of a and λ on the effect of α are ambiguous.

If C''' is unequal zero, then the signs of the cross-partial derivatives are, in general, ambiquous.

The corollary shows that higher responsiveness to quality reduces how responsive quality is to changes in the number of providers (in equilibrium of the quality subgame). The effect is the opposite if the patients with inaccurate perceptions tend to overreact to quality and if the share of patients with accurate perceptions increases, that is when λ becomes higher. Moreover, the corollary reveals that higher responsiveness to quality intensifies the effect of the prospective payment on equilibrium quality: higher a increases how responsive q is to changes in p on the margin. The effect is the opposite if the patients with inaccurate tend to overreact to quality and if the share of patients with accurate perceptions increases.

4.3.2 Entry

Consider a subgame in Stage 2, defined by a payment scheme $R = (p, \alpha)$. If n providers enter, the continuation equilibrium profits, given symmetric equilibrium qualities $q^*(n; R)$ in Stage 3, are

$$\left[p - (1-\alpha)cq^*\left(n; R\right)\right]\frac{1}{n} - C\left(q^*(n; R)\right) - K, \tag{8}$$

in which $\frac{1}{n}$ is the demand in the symmetric equilibrium, and K is the fixed entry cost. $q^*\left(n; R\right)$ is the symmetric equilibrium quality from Proposition 1, in which I have emphasized that q^* depends on the number of providers, n, and on the regulated reimbursement scheme, R.

The profits are positive in a monopoly (by assumption). Then for $n = 1$, equation (8) becomes $p - (1-\alpha)cq\left(1\right) - C\left(q\left(1\right)\right) - K > 0$. Differentiating (8) with respect to n and using Corollary 1 gives:

$$-\frac{p - (1-\alpha)cq^*(n; R)}{n^2} - \left[\frac{(1-\alpha)c}{n} + C'\left(q^*(n; R)\right)\right]\frac{\partial q^*(n; R)}{\partial n} < 0. \tag{9}$$

Equation (9) says that the profits monotonically decrease in n. An increase in the number of providers has a direct negative effect on the profits of providers because it leads to lower market shares. This is the first part of (9). Due to Corollary 1 $\frac{\partial q^*}{\partial n} > 0$, which means greater entry leads to higher quality. This increased quality, in turn, increases the costs of quality production. This is the second term in (9).

The free entry equilibrium number of providers n^* is determined from the zero profit condition for the providers. The number of health care providers entering the market is an integer.[13] Because profits are positive when n is small and

---

13 The market coverage condition requires that $\bar{s} > -q^* + \frac{\tau}{2n^*} \equiv \widehat{s}$, where q^* is given in Proposition 1 and n^* is the equilibrium number of providers in Stage 2. Therefore, if $\bar{s}$ is large enough, the market coverage condition is satisfied. Note that neither q^* nor n^* depend on $\bar{s}$.

profits are negative for large n, there is a unique solution n^* for which the profits are zero. Formally

$$\pi(n^*; y) \equiv \frac{p - (1-\alpha)cq^*(n^*; R)}{n^*} - C\left(q^*(n^*; R)\right) - K = 0. \tag{10}$$

where $y = (z, K)$.

Proposition 2 *In subgame $R = (p, \alpha)$, the equilibrium number of health care providers n^* is determined implicitly by equation (10).*

Equation (10) defines implicitly n^* as a function of parameters $y = (z, K)$.

Next I present the comparative statics of n^*. I begin with the comparative statistics with respect to competition parameters. I have summarized the results in the following corollary.

Corollary 5 *n^* decreases in the entry cost K and increases in τ.*

Corollary 5 shows what happens if the competition environment changes. When the entry cost is lower, a larger number of providers enter, and the market becomes more competitive. With a larger number of providers, each patient can choose a service closer to their preferred one. A decrease in the transportation cost decreases the degree of differentiation. Thus, post-entry competition increases, and provider entry is reduced.

The following shows how equilibrium entry is affected by payment scheme regulation.

Corollary 6 *The effect of p and α on n^* is, in general, ambiguous. The effects of both payment instruments are positive if the indirect cost effect, which a higher equilibrium quality causes, is smaller than the direct profitability effect.*

Increasing cost sharing, α, has two effects. First, it decreases the per-unit variable costs. This direct effect increases the per-unit profit margin and therefore boosts entry. Second, increasing cost sharing leads to higher quality in equilibrium and thus to increased variable costs. This indirect effect decreases entry. The total effect is hence ambiguous, unless further assumptions are imposed. If the direct profitability effect dominates, the total effect of increasing cost sharing is that more providers enter.

The effect of prospective payment is also ambiguous in general. Like the cost sharing parameter, the prospective payment affects entry through two channels. First, the direct effect of higher prospective payment on the per-unit profit margin. This effect increases entry. The second effect is the indirect effect that comes about via increased equilibrium quality. Higher quality results in higher costs, which decreases entry. If the direct profitability effect dominates, the total effect of higher prospective payment is that more providers enter.

To sign the derivatives conclusively requires imposing more structure in the model. One possibility would be to assume that the cost of providing higher quality is very convex, that is C'' is very large. Then the direct effect dominates, and entry increases with cost sharing and prospective payment.

The following presents how patient misperception affects entry.

Corollary 7 *n^* decreases in a. If patients overreact to quality, n^* increases in λ. The effect is the opposite if patients underreact to quality.*

In particular, increasing the responsiveness to quality leads to an increase in the quality elasticity of demand of those patients who perceive quality imperfectly. As discussed earlier, this increases the overall quality elasticity of demand. More elastic demand leads to fiercer competition in qualities and thus to higher variable costs. The higher costs, in turn, lead to lower profits and reduced entry. The effect of more information available to patients about the provider quality, that is λ is higher, depends on whether patients overreact or underreact to quality. If patients overreact to perceived quality, an increase in the share of patients with accurate perception increases the equilibrium number of providers. The reason is the indirect effect, which results in less elastic demand, lower qualities (and the associated variable costs) and thus to increased profitability and entry. The indirect effect works in the opposite way if patients underreact to perceived quality.

4.4 Payment scheme regulation

In this section, I analyze the optimal payment regulation scheme. I begin by characterizing the first best. Then I consider the first regulatory framework with contractible entry and provider payment. Last, I consider the second regulatory framework when entry is noncontractible and the regulator uses a mixed payment scheme.

4.4.1 First best

Consider a utilitarian social welfare function where the social planner gives equal weight to the providers and patients.[14] The first best consists of a vector of qualities Q and a number of providers n, that is (Q, n), which maximizes the following social welfare

$$W(Q,n) = \left\{ \bar{s} + q - \tau \left[2n \int_0^{1/2n} x dx \right] \right\} + n\pi - \left\{ n \left[(p + (c - (1-\alpha)c)q) \frac{1}{n} \right] \right\}. \tag{11}$$

The first term in curly brackets gives the surplus of the patients, the second term is the total surplus (profits) for the providers, and the last term in curly brackets denotes the regulatory costs of health care provision. Making use of symmetry, social welfare becomes:[15]

$$W(Q,n) = \bar{s} + q - \frac{\tau}{4n} - nC(q) - nK - cq. \tag{12}$$

14 Equal weight given to patients and providers means that the costs can be paid lump sum by either providers or patients without changing the level of social surplus.

15 See Appendix 4.A.3 for the details of the derivation of this welfare function.

The social planner takes into account social benefits and costs of quality and minimizes the sum of the fixed costs and the patients' transportation costs. Differentiating (12) with respect to q and n and setting these first-order derivatives to zero gives:

$$O^1(q,n;c,\tau,K) \equiv (1-c)\frac{1}{n} - C'(q) = 0 \tag{13}$$

$$O^2(q,n;c,\tau,K) \equiv \frac{\tau}{4n^2} - C(q) - K = 0. \tag{14}$$

First-order conditions (13) and (14) implicitly define q^{FB} and n^{FB} as functions of (c,τ,K)

Proposition 3 *First best q^{FB} and n^{FB} are determined implicitly by equations (13) and (14).*

This proposition shows that the first best quality and the number of providers depend on the cost of producing quality (c and the parameters of the $C(.)$-function), the degree of differentiation in the market (τ), and the fixed cost of entry (K).

After applying the implicit function theorem to (13) and (14), I obtain the following two corollaries. The following gives how the first best is affected by the variable costs and the competition parameters.

Corollary 8 *q^{FB} decreases in c and τ, but increases in K. n^{FB} increases in c and τ, but decreases in K.*

The intuition for the effects of the fixed entry costs, K, is easy to see. Higher entry cost means that the social cost of an additional firm is higher. This reduces the first best number of providers. This, in turn, reduces the social marginal cost of increasing quality. The first best quality increases therefore with the entry costs.

The first best quality decreases in the variable cost of producing quality, c, because it increases the marginal social cost of providing higher quality to the patients. The first best number of providers increases in the variable cost of producing quality. This indirect effect arises from the lower fixed cost of producing quality, $C(q)$. The lower quality means that the social cost of an additional firm is lower, leading to a higher number of providers in the first best.

When the degree of differentiation increases, that is the transportation cost τ increases, the marginal social benefit of having one more firm in the market increases. The reason for this is that having more providers reduces the transportation costs that patients face. Hence, the first best number of providers increases with the transportation costs. However, having more providers increases the social cost of providing quality because of the fixed, volume-independent, component of the quality cost function. Hence, higher transportation costs result in a lower first best quality.

The following summarizes how the first best is affected by misperception parameters.

Corollary 9 *q^{FB} and n^{FB} are independent of a and λ.*

The regulator chooses q^{FB} and n^{FB} to maximize social welfare, and social welfare is based on experienced true benefit, not on patients' decision utility. Therefore, regulator does not take a or λ into account in the welfare calculations, and q^{FB} and n^{FB} are unaffected by patient misperception.

4.4.2 Equilibrium qualities and the number of providers when the number of providers is contractible

In this subsection, I study the optimal payment scheme regulation in the first regulatory framework. In this framework the regulator can control the number of providers in the market and uses a single regulatory instrument, a prospective payment. This kind of situation is of interest because entry regulation is common in many health care markets, such as in medical licensing and in the pharmacy market. The following proposition states that if the first best number of providers is contractible, the first-best quality level is achieved by using prospective payment only.

Proposition 4 *When the number of providers is contractible, the regulator can implement the first best quality and the number of providers by granting licenses for n^{FB} providers and by setting the following prospective payment*

$$\widehat{p} = cq^{FB} + \frac{\tau}{\kappa n^{FB}} \tag{15}$$

where (q^{FB}, n^{FB}) solve (13) and (14).

This proposition shows that when the number of providers is contractible, the regulator can use entry licensing and a prospective payment to achieve the first-best outcome. This is because the regulator chooses the number of entry licenses so that it satisfies equation (14) for any value of quality. This is achieved by choosing the level of prospective payment so that the providers implement the quality level that satisfies (13) for any number of providers in the market. If the number of entry licenses and the prospective payment are such that equations (13) and (14) hold simultaneously, the first best ensues.

The result is similar to the result originally presented in Ma (1994) and later extended by Brekke et al. (2006) and Bardey et al. (2012). In particular, Bardey et al. (2012) show that when the regulator can control the location of a health care provider (in other words, the level of horizontal product differentiation), it can implement the first-best quality by a certain prospective payment. As far as I am aware, the prior literature has not acknowledged the possibility that entry licensing combined with a prospective payment can achieve the first-best quality and market structure.

Corollary 10 *If the prospective payment satisfies (15) and the number of entry licenses is granted so that the first-best is obtained, a higher K increases $\widehat{p}$. However, the effect of τ and c on $\widehat{p}$ is ambiguous.*

The intuition for this corollary is the following. Note that higher entry cost, K, has no direct effect on the prospective payment, $\widehat{p}$. Corollary 10 gives the indirect effects: higher entry reduces the first best number of providers but increases the first best quality. Equation (15) shows that when quality increases or the number of providers decrease, the prospective payment increases. Thus, even when the regulator does not directly reimburse the fixed costs of providers, it does so indirectly.

The corollary also shows that the effect of the transportation costs and the quality-related variable costs on $\widehat{p}$ is ambiguous. This is because these parameters have both a positive direct effect on $\widehat{p}$ and a negative indirect effect on $\widehat{p}$. Regarding the positive direct effects, note that when the variable cost of quality provision increases, the prospective payment, $\widehat{p}$, becomes higher. The intuition for this direct effect is that an increase in the variable cost would decrease the equilibrium quality chosen by providers; this arises when the implicit function theorem is applied to (7). To be able to induce the first best quality the regulator increases the prospective payment. This offsets the increase in variable costs. An increase in the transportation costs weakens competition in qualities (see Corollary 1). To offset this decrease in competition, the regulator provides higher prospective payment. This explains the positive direct effect of transportation costs.

Thus, the total effects of τ and c on $\widehat{p}$ are ambiguous, because there is the negative indirect effects that follow partly from Corollary 8 and from how q^{FB} and n^{FB} enter (15).

The following summarizes how $\widehat{p}$ depends on misperception parameters.

Corollary 11 *If the prospective payment satisfies (15) and the number of entry licenses is granted so that the first-best is obtained, higher a decreases $\widehat{p}$. Moreover, if some patients overreact to quality $(a > 1)$, increase in λ increases $\widehat{p}$. If some patients underreact to quality $(a < 1)$, the effect of λ is the opposite.*

Intuition for this result is the following. Recall that the first best quality and number of providers are not affected by the misperception parameters by Corollary 9. Thus, there are no indirect effects. The direct effects follow because the prospective payment is inversely related to κ. Thus, when the demand becomes more responsive to quality, the providers choose higher qualities in equilibrium. The optimal response of the regulator is to decrease the prospective payment. A similar reasoning applies to the increase in λ when the patients with inaccurate perceptions overreact to the quality.

4.4.3 Equilbrium qualities and the number of providers when the number of providers is not contractible

In this subsection, the number of providers is not contractible, and the regulator uses provider payment and cost reimbrusement as its payment scheme. The regulator's problem is to choose a payment scheme $R = (p, \alpha)$ that maximizes a utilitarian social welfare function. Given the continuation equilibrium number of providers n^* and qualities q^*, the parameters p and α are solutions to the following problem:

$$\max_{R=(p,\alpha)} \bar{s} + q^*\left(n^*(R), R\right) - \frac{\tau}{4n^*(R)} - n^*(R)C\left(q^*\left(n^*(R), R\right)\right) \\ - Kn^*(R) - cq^*\left(n^*(R), R\right) \quad (16)$$

subject to $\alpha \geq 0$ and $\alpha \leq 1$. To characterize the regulator's choices of $R = (p, \alpha)$ and the associated market equilibrium, I form a Lagrangian function for the optimization problem (16) as follows

$$L = \bar{s} + q^*\left(n^*(R), R\right) - \frac{\tau}{4n^*(R)} - n^*(R)C\left(q^*\left(n^*(R), R\right)\right) - Kn^*(R) \\ - cq^*\left(n^*(R), R\right) - \nu\left[(1-\alpha) - 1\right] + \mu\left[1 - \alpha\right] \quad (17)$$

I study three possible regimes. These regimes depend on whether and which of the two constraints bind. In Regime A, neither of the constraints bind, $\nu = \mu = 0$. In this regime, the regulator uses a mixed payment instrument that consists of a prospective payment and a partial cost reimbursement such that $(1-\alpha)c$ belongs to the open interval $(0, c)$. In Regime B, $\alpha < 1$ -constraint binds, that is $\nu > 0$ and $\mu = 0$. In this regime, the regulator offers full cost reimbursement, and $(1-\alpha)c$ is constrained to be equal to zero. Lastly, in Regime C, $\alpha = 0$ constraint binds, that is $\nu = 0$ and $\mu > 0$. In this regime, the regulator uses pure prospective payment, does not use cost reimbursement, and $(1-\alpha)c$ is constrained to be equal to c.[16]

I differentiate the Lagrangian in (17) with respect to p and α to obtain the following first-order conditions:

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R), R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}p} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R), R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}p} = 0 \quad (18)$$

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R), R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R), R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}\alpha} \\ + \nu - \mu = 0, \quad (19)$$

in which $\frac{\mathrm{d}q^*}{\mathrm{d}p} = \frac{\partial q^*}{\partial n^*}\frac{\partial n^*}{\partial p} + \frac{\partial q^*}{\partial p}$ and $\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} = \frac{\partial q^*}{\partial n^*}\frac{\partial n^*}{\partial \alpha} + \frac{\partial q^*}{\partial \alpha}$.

Unfortunately, characterizing when each of the three regimes prevails turns out to be analytically intractable. However, I can use (18) and (19) to state the following result.

Proposition 5 *If Regime A prevails, and if there is an interior solution to the problem defined by the Lagrangian in (17), the first best can be implemented with the following mixed payment scheme:*

$$p^A = \left\{\frac{\kappa(\lambda, a)(n^{FB})^2\left[C\left(q^{FB}\right) + K\right]}{\tau} - C'(q^{FB})n^{FB}\right\}q^{FB} + \left[C\left(q^{FB}\right) + K\right]n^{FB}$$

---

16 For L I also use (implicitly) the earlier assumption of non-negative prospective payment.

and

$$\alpha^A = 1 - \frac{1}{c}\left\{\frac{\kappa(n^{FB})^2\left[C\left(q^{FB}\right)+K\right]}{\tau} - C'(q^{FB})n^{FB}\right\},$$

where (q^{FB}, n^{FB}) *are as in Proposition 3.*

Proposition 5 shows that if the two constraints on the cost reimbursement do not bind, the first-best can be implemented.

This has several implications. First, the comparative statics of p^A and α^A with many respect to many model parameters cannot be signed without further restrictions. The following summarizes this and how p^A and α^A depend on misperception parameters.

Corollary 12 *If Regime A prevails and* p^A *and* α^A *are set such as in Proposition 5, higher* a *increases* p^A *and decreases* α^A*. Moreover, if some patients overreact to quality* $(a > 1)$*, increase in* λ *decreases* p^A *and increases* α^A*. If some patients underreact to quality* $(a < 1)$*, the effect of* λ *is the opposite. The effects of* K*,* τ*, and* c *on* p^A *and* α^A *are ambiguous.*

Second, note that I require that $0 < \alpha^A < 1$ holds in the interior solution of Regime A. Unfortunately, determining when this requirement holds is not possible without further assumptions. However, if the model parameters are such that Regime A cannot prevail, I can state the following:

Proposition 6 *If Regime B prevails or if Regime C prevails, then the first best cannot be implemented.*

This result shows that when the overcompensation of the costs is not allowed ($\alpha < 1$), and when negative compensation is not allowed either ($0 < \alpha$), the regulator does not have enough instruments to achieve the first best. This result is similar to Bardey et al. (2012) who find that in health care markets with horizontal and vertical differentiation a mixed reimbursement scheme is welfare improving. In addition to their efficiency and inefficiency results, Bardey et al. (2012) are able to pin down exactly the regime where this happens; that is if the an equilibrium allocation induced by using prospective payment only is characterized by too low quality (vertical differentiation) and overspecialization (horizontal differentiation).

4.5 Policy discussion

My results have two main policy implications. First, my results show that entry licensing and prospective payment implement the first best equilibrium outcome. When the number of providers in the market is not contractible, the regulator may be able to implement the first best by using the prospective payment and the cost reimbursement. The first best cannot be obtained, if the economic environment is such that it calls for either full cost sharing or no cost sharing.

My second policy implication relates to if the regulator were to reduce quality misperception by providing patients with more information about the provider quality. An example of such a policy would be use of quality report cards and other similar ways of disclosing quality information to the patients. This information provision would increase the share of patients with accurate quality information. The effects of such information provision are however hard to predict: the previous literature has shown that quality reporting might lead to selection/cream skimming issues, or other strategic responses from the supply side such as providers putting more effort on the measured quality measures which may not be the ones that maximize patient health and welfare. In addition to these concerns, my results suggest that the effects of providing more information about the provider quality may have different direct and indirect effects. The direct and indirect effects also depend on whether patients underreact or overreact to quality.

4.6 Conclusions

I have studied the regulation of health care payment schemes for health care providers entry decisions and quality choices when some patients have inaccurate quality perceptions that affect their decision utilities. I have extended the model of Salop (1979) as follows: i) instead of a price, providers choose a quality of a medical good or a service that they offer; ii) the regulator covers all medical expenses for the patient, and the only cost patients face is a transportation cost to a provider location: iii) the welfare maximizing regulator chooses how the payment scheme is regulated: and iv) some patients may overreact or underreact to the perceived health care quality. To my knowledge, this is the first model that combines provider entry and quality choices, patient quality misperception, and payment regulation in the same framework.

I found that if the regulator can control the number of health care providers, the first best quality may be obtained with a combination of entry regulation and a pure prospective payment. However, if this were not the case, a mixed payment scheme would be needed to obtain efficiency. Regarding quality misperception, the more responsive patients were to the perceived quality, the higher the equilibrium quality. It also reduced entry. Quality misperception weakened the effect of the number of providers and strengthened the effect of prospective payment on equilibrium quality.

The model studied in this paper rules out mixed strategies which would, if allowed, in equilibrium be used. Even in the context of pure strategies, the equilibrium requires some coordination or sequencing of decision, and identity of the entrants is not determined. While Salops model is useful in studying the effects of changes in market structure, the model when studying entry, because a sequential entry would force all incumbent firms to relocate and a simultaneous entry would require coordination in terms of location. These problems are, however, shared by virtually all other papers that use Salops model to study entry.

The analysis here could be extended in several ways. Understanding the mechanisms for the inefficient equilibrium, that is whether there is over or under entry and whether the quality is too high or too low, is an obvious next step. Also whether patient misperception alleviates or strengthens the effects of competition and payment scheme regulation should be studied further. A potential extension of this model would be to consider a case where the firms prices are restricted to be same but not necessarily to zero.

Last, analyzing how providers choose their locations after their entry decision, rather than having an auctioneer choose the particular location pattern in addition to entry deterrence are left for future analysis too.

4.A Appendix: Proofs and derivations

This appendix contains the proofs of the results reported in the main text. Appendix 4.A.1 collects the proofs of the results in Subsection 4.3.1 Appendix A.2 collects the proofs of the results in Subsection 4.3.2 Appendix 4.A.3 collects the proofs of the results in Subsection 4.4.1 Appendix 4.A.4 collects the proofs of the results in Subsection 4.4.2 Appendix 4.A.5 collects the proofs of the results in Subsection 4.4.3 Lastly, Appendix provides an example with more specific assumptions on the functional form of the production costs.

4.A.1 Proofs of the results in Section 4.3.1

Proof of Proposition 1

In this proof I characterize the symmetric Nash equilibrium qualities of the game. I begin the proof by solving the first-order condition. Then I show that the second-order condition is negative, and therefore the first-order condition is sufficient for profit maximization. Last, I use the assumption of symmetric Nash equilibrium to simplify the first-order condition. This simplified first-order condition defines q as an implicit function of the exogenous parameters.

Provider i's maximization problem is to choose its quality q_i to maximize its profits in (5), given the total demand (4) and the other providers' qualities q_{-i}. Equilibrium qualities $Q^* = \left(q_1^*, ..., q_i^*, ..., q_n^*\right)$ are best responses against other providers' qualities.

I obtain the first-order derivative by differentiating (5) with respect to q_i. Setting the first-order derivative to zero gives the following first-order condition:

$$[p - (1-\alpha)cq_i] \frac{\partial D_i\left(q_i, q_-i\right)}{\partial q_i} - (1-\alpha)cD_i\left(q_i, q_-i\right) - C'\left(q_i\right) = 0. \tag{20}$$

I find the second-order condition and its sign next. Differentiating (20) with respect to q_i and using $\frac{\partial D_i(q_i, q_-i)}{\partial q_i} = \frac{\lambda}{\tau} + \frac{(1-\lambda)a}{\tau} = \frac{\kappa(a,\lambda)}{\tau}$, the second-order condition becomes

$$\frac{\partial^2 \pi_i}{\partial q_i \partial q_i} = -\frac{(1-\alpha)c\kappa\left(a,\lambda\right)}{\tau} - (1-\alpha)c \underbrace{\frac{\partial D_i\left(q_i, q_-i\right)}{\partial q_i}}_{=\frac{\kappa(a,\lambda)}{\tau}} - C''\left(q_i\right). \tag{21}$$

Because $C''\left(q_i\right) > 0$, the sign of (21) is negative, and the first-order condition in (20) is sufficient for profit maximization.

I focus on a symmetric equilibrium in which providers are located equidistantly. Imposing symmetry $q = q_i = q_{-i}$, the total demand in (4) becomes $\frac{1}{n}$. I use this to simplify the first-order condition (20) so that it becomes

$$[p - (1-\alpha)cq] \frac{\kappa\left(\lambda, a\right)}{\tau} - \frac{(1-\alpha)c}{n} - C'\left(q\right) = 0. \tag{22}$$

Last, I use (22) to construct the following:

$$G(q,z) \equiv [p-(1-\alpha)cq]\,\frac{\kappa\,(\lambda,a)}{\tau} - \frac{(1-\alpha)c}{n} - C'\,(q) = 0, \tag{23}$$

which defines q as an implicit function of $z = (p,\alpha,c,n,\tau,a,\lambda)$. Equilibrium quality q^* is the value of q that solves function G implicitly as a function of parameters $z = (p,\alpha,c,n,\tau,a,\lambda)$, that is the symmetric Nash equilibrium quality q^* is determined implicitly by $G(q^*;z) = 0$, where $z = (p,\alpha,c,n,\tau,a,\lambda)$. Note that using the assumption of $p > \frac{c\tau}{\kappa(\lambda,a)}$, $G(0,z) = \frac{\kappa p}{\tau} - \frac{(1-\alpha)c}{n} > 0$ for all $n \geq 1$ and all $0 \leq \alpha \leq 1$. Then $p-(1-\alpha)cq^* > 0$ because $p > \frac{c\tau}{\kappa}$, which implies that the providers choose a strictly positive quality in the symmetric equilibrium. This completes the proof.

Proof of Corollary 1

In this proof I find the comparative statics of the equilibrium quality q^* with respect to competition parameters n and τ by using the implicit function theorem around q^*. I have omitted the arguments λ and a from κ to clarify the notation.

I begin by obtaining the comparative statics of q^* with respect to the number of providers n. Using the implicit function theorem around q^* gives

$$\frac{\partial q^*}{\partial n} = -\frac{\frac{\partial G}{\partial n}}{\frac{\partial G}{\partial q}} = \frac{\dfrac{(1-\alpha)c}{n^2}}{(1-\alpha)c\frac{\kappa}{\tau} + C''(q^*)} > 0,$$

in which (21) is used for $\frac{\partial G}{\partial q}$.

Then, I obtain the comparative statics of q^* with respect to the transportation cost τ. Using the implicit function theorem around q^* gives

$$\frac{\partial q^*}{\partial \tau} = -\frac{\frac{\partial G}{\partial \tau}}{\frac{\partial G}{\partial q}} = \frac{-\left[p-(1-\alpha)cq^*\right]\dfrac{\kappa}{\tau^2}}{(1-\alpha)c\frac{\kappa}{\tau} + C''(q^*)} < 0,$$

in which (21) is used for $\frac{\partial G}{\partial q}$ and because $p-(1-\alpha)cq^* > 0$ (see proof of Proposition 1). This completes the proof.

Proof of Corollary 2

In this proof I find the comparative statics of quality q^* with respect to regulation parameters p and α by using the implicit function theorem around q^*. I have omitted the arguments λ and a from κ to clarify the notation.

I begin by obtaining the comparative statics of q^* with respect to prospective payment p. Using the implicit function theorem around q^* gives

$$\frac{\partial q^*}{\partial p} = -\frac{\frac{\partial G}{\partial p}}{\frac{\partial G}{\partial q}} = \frac{\dfrac{\kappa}{\tau}}{(1-\alpha)c\frac{\kappa}{\tau} + C''(q^*)} > 0,$$

in which (21) is used for $\frac{\partial G}{\partial q}$.

Then, I obtain the comparative statics of q^* with respect to cost reimbursement α. Using the implicit function theorem around q^* gives

$$\frac{\partial q^*}{\partial \alpha} = -\frac{\frac{\partial G}{\partial \alpha}}{\frac{\partial G}{\partial q}} = \frac{c\left(\frac{q^*\kappa}{\tau} + \frac{1}{n}\right)}{(1-\alpha)c\frac{\kappa}{\tau} + C''(q^*)} > 0,$$

in which (21) is used for $\frac{\partial G}{\partial q}$. This completes the proof.

Proof of Corollary 3

In this proof I find the comparative statics of q^* with respect to misperception parameters a and λ using the implicit function theorem around q^*. The sign of $\frac{\partial q^*}{\partial \lambda}$ depends on whether $a < 1$ or $a > 1$. Therefore, the proof for $\frac{\partial q^*}{\partial \lambda}$ consists of two parts. Note, that $\frac{\partial \kappa(\lambda,a)}{\partial a} = 1 - \lambda$ and $\frac{\partial \kappa(\lambda,a)}{\partial \lambda} = 1 - a$.

I begin by obtaining the comparative statics of q^* with respect to patients' responsiveness on quality a. Using the implicit function theorem around q^* gives

$$\frac{\partial q^*}{\partial a} = -\frac{\frac{\partial G}{\partial a}}{\frac{\partial G}{\partial q}} = \frac{[p - (1-\alpha)cq^*]\,\frac{(1-\lambda)}{\tau}}{(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)} > 0$$

in which (21) is used for $\frac{\partial G}{\partial q}$ and because $p - (1-\alpha)cq^* > 0$ (see proof of Proposition 1).

I obtain the comparative statics of q^* with respect to the share of patients with accurate perceptions λ next. I do this in two parts.

First, let $a > 1$. Using the implicit function theorem around q^* gives

$$\frac{\partial q^*}{\partial \lambda} = -\frac{\frac{\partial G}{\partial \lambda}}{\frac{\partial G}{\partial q}} = \frac{[p - (1-\alpha)cq^*]\,\frac{(1-a)}{\tau}}{(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)} < 0. \tag{24}$$

in which (21) is used for $\frac{\partial G}{\partial q}$ and because $p - (1-\alpha)cq^* > 0$ (see proof of Proposition 1).

Last, let $a < 1$. Using the implicit function theorem around q^* gives

$$\frac{\partial q^*}{\partial \lambda} = -\frac{\frac{\partial G}{\partial \lambda}}{\frac{\partial G}{\partial q}} = \frac{[p - (1-\alpha)cq^*]\,\frac{(1-a)}{\tau}}{(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)} > 0. \tag{25}$$

in which (21) is used for $\frac{\partial G}{\partial q}$ and because $p - (1-\alpha)cq^* > 0$ (see proof of Proposition 1). This completes the proof.

Proof of Corollary 4

In this proof I study whether patient misperception alleviates or strengthens the effects of the competition and payment scheme regulation. I obtain the cross-partial derivatives by differentiating $\frac{\partial q^*}{\partial n}$, $\frac{\partial q^*}{\partial \tau}$, $\frac{\partial q^*}{\partial p}$, and $\frac{\partial q^*}{\partial \alpha}$ with respect to a and

λ. Because the signs of the cross-partial derivatives with respect to λ depend on whether $a < 1$ or $a > 1$, I will consider these cases separately.

I begin with the cross-partial derivatives with respect to a. Differentiating $\frac{\partial q^*}{\partial n}$ from the proof of Corollary 1 with respect to a gives

$$\frac{\partial^2 q^*}{\partial n \partial a} = \frac{-\frac{(1-\alpha)c}{n^2}\left[(1-\alpha)c\frac{(1-\lambda)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial a}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.$$

Because $\frac{\partial q^*}{\partial a} > 0$ the cross-partial derivative is negative if $C'''(q^*) = 0$. Otherwise the sign is ambiguous.

Differentiating $\frac{\partial q^*}{\partial \tau}$ from the proof of Corollary 1 with respect to a gives

$$\begin{aligned}
\frac{\partial^2 q^*}{\partial \tau \partial a} =& \Bigg\{ \left\{ \left[(1-\alpha)c\frac{\partial q^*}{\partial a}\right]\frac{\kappa(\lambda,a)}{\tau^2} - [p-(1-\alpha)cq^*]\frac{1-\lambda}{\tau^2} \right\} \left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] \\
&+ [p-(1-\alpha)cq^*]\frac{\kappa(\lambda,a)}{\tau^2}\left[(1-\alpha)c\frac{1-\lambda}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial a}\right] \Bigg\} \\
&\frac{1}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2} \\
=& \Bigg\{ (1-\alpha)^2c^2\frac{\partial q^*}{\partial a}\frac{\kappa(\lambda,a)^2}{\tau^3} + (1-\alpha)c\frac{\partial q^*}{\partial a}C''(q^*) - [p-(1-\alpha)cq^*]\frac{1-\lambda}{\tau^2}C''(q^*) \\
&+ [p-(1-\alpha)cq^*]\frac{\kappa(\lambda,a)}{\tau^2}C'''(q^*)\frac{\partial q^*}{\partial a} \Bigg\} \frac{1}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.
\end{aligned}$$

The sign is ambiguous.

Differentiating $\frac{\partial q^*}{\partial p}$ from the proof of Corollary 2 with respect to a gives

$$\begin{aligned}
\frac{\partial^2 q^*}{\partial p \partial a} &= \frac{\frac{1-\lambda}{\tau}\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] - \frac{\kappa}{\tau}\left[(1-\alpha)c\frac{(1-\lambda)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial a}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2} \\
&= \frac{\frac{1-\lambda}{\tau}C''(q^*) - \frac{\kappa}{\tau}C'''(q^*)\frac{\partial q^*}{\partial a}}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.
\end{aligned}$$

Because $\frac{\partial q^*}{\partial a} > 0$ the cross-partial derivative is positive if $C'''(q^*) = 0$. Otherwise the sign is ambiguous.

Differentiating $\frac{\partial q^*}{\partial \alpha}$ from the proof of Corollary 2 with respect to a gives

$$\frac{\partial^2 q^*}{\partial \alpha \partial a} = \frac{c\left[\frac{\frac{\partial q^*}{\partial a}\kappa + q^*(1-\lambda)}{\tau}\right]\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] - c\left(\frac{q^*\kappa}{\tau} + \frac{1}{n}\right)\left[(1-\alpha)c\frac{(1-\lambda)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial a}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.$$

The sign is ambiguous.

Consider the cross-partial derivatives with respect to the share of patients with accurate quality perception λ next.

First, let $a > 1$. I differentiate $\frac{\partial q^*}{\partial n}$ from the proof of Corollary 1 with respect to λ and obtain

$$\frac{\partial^2 q^*}{\partial n \partial \lambda} = \frac{-\frac{(1-\alpha)c}{n^2}\left[(1-\alpha)c\frac{(1-a)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.$$

Because $\frac{\partial q^*}{\partial \lambda} < 0$ the cross-partial derivative is positive if $C'''(q^*) = 0$. Otherwise the sign is ambiguous.

Differentiating $\frac{\partial q^*}{\partial \tau}$ from the proof of Corollary 1 with respect to λ gives

$$\begin{aligned}\frac{\partial^2 q^*}{\partial \tau \partial \lambda} = &\left\{ \left\{ \left[(1-\alpha)c\frac{\partial q^*}{\partial \lambda}\right] \frac{\kappa(\lambda,a)}{\tau^2} - [p-(1-\alpha)cq^*]\frac{1-a}{\tau^2} \right\} \left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] \right. \\ &\left. + [p-(1-\alpha)cq^*]\frac{\kappa(\lambda,a)}{\tau^2}\left[(1-\alpha)c\frac{1-a}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right] \right\} \\ &\frac{1}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2} \\ &= \left\{ (1-\alpha)^2 c^2 \frac{\partial q^*}{\partial \lambda}\frac{\kappa(\lambda,a)^2}{\tau^3} + (1-\alpha)c\frac{\partial q^*}{\partial \lambda}C''(q^*) - [p-(1-\alpha)cq^*]\frac{1-a}{\tau^2}C''(q^*) \right. \\ &\left. + [p-(1-\alpha)cq^*]\frac{\kappa(\lambda,a)}{\tau^2}C'''(q^*)\frac{\partial q^*}{\partial \lambda} \right\} \frac{1}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.\end{aligned}$$

Because $a > 1$ and $\frac{\partial q^*}{\partial \lambda} < 0$ the sign of the cross-partial derivative is ambiguous.

Differentiating $\frac{\partial q^*}{\partial p}$ from the proof of Corollary 2 with respect to λ gives

$$\begin{aligned}\frac{\partial^2 q^*}{\partial p \partial \lambda} &= \frac{\frac{1-a}{\tau}\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] - \frac{\kappa}{\tau}\left[(1-\alpha)c\frac{(1-a)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2} \\ &= \frac{\frac{1-a}{\tau}C''(q^*) - \frac{\kappa}{\tau}C'''(q^*)\frac{\partial q^*}{\partial \lambda}}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.\end{aligned}$$

Because $a > 1$ and $\frac{\partial q^*}{\partial \lambda} < 0$ the cross-partial derivative is negative if $C'''(q^*) = 0$. Otherwise the sign is ambiguous.

Differentiating $\frac{\partial q^*}{\partial \alpha}$ from the proof of Corollary 2 with respect to λ gives

$$\frac{\partial^2 q^*}{\partial \alpha \partial \lambda} = \frac{c\left[\frac{\frac{\partial q^*}{\partial \lambda}\kappa + q^*(1-a)}{\tau}\right]\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] - c\left(\frac{q^*\kappa}{\tau} + \frac{1}{n}\right)\left[(1-\alpha)c\frac{(1-a)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.$$

Because $a > 1$ and $\frac{\partial q^*}{\partial \lambda} < 0$ the sign of the cross-partial derivative is ambiguous.

Last, let $a < 1$. I differentiate $\frac{\partial q^*}{\partial n}$ from the proof of Corollary 1 with respect to λ and obtain

$$\frac{\partial^2 q^*}{\partial n \partial \lambda} = \frac{-\frac{(1-\alpha)c}{n^2}\left[(1-\alpha)c\frac{(1-a)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.$$

Because $\frac{\partial q^*}{\partial \lambda} > 0$ the cross-partial derivative is negative if $C'''(q^*) = 0$. Otherwise the sign is ambiguous.

Differentiating $\frac{\partial q^*}{\partial \tau}$ from the proof of Corollary 1 with respect to λ gives

$$\begin{aligned}
\frac{\partial^2 q^*}{\partial \tau \partial \lambda} = &\left\{ \left\{ \left[(1-\alpha)c\frac{\partial q^*}{\partial \lambda}\right]\frac{\kappa(\lambda,a)}{\tau^2} - [p-(1-\alpha)cq^*]\frac{1-a}{\tau^2}\right\}\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] \right.\\
&\left. + [p-(1-\alpha)cq^*]\frac{\kappa(\lambda,a)}{\tau^2}\left[(1-\alpha)c\frac{1-\lambda}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right]\right\}\\
&\frac{1}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}\\
= &\left\{(1-\alpha)^2c^2\frac{\partial q^*}{\partial \lambda}\frac{\kappa(\lambda,a)^2}{\tau^3} + (1-\alpha)c\frac{\partial q^*}{\partial \lambda}C''(q^*) - [p-(1-\alpha)cq^*]\frac{1-a}{\tau^2}C''(q^*)\right.\\
&\left. + [p-(1-\alpha)cq^*]\frac{\kappa(\lambda,a)}{\tau^2}C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right\}\frac{1}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.
\end{aligned}$$

Because $a < 1$ and $\frac{\partial q^*}{\partial \lambda} > 0$ the sign of the cross-partial derivative is ambiguous.

Differentiating $\frac{\partial q^*}{\partial p}$ from the proof of Corollary 2 with respect to λ gives

$$\begin{aligned}
\frac{\partial^2 q^*}{\partial p \partial \lambda} &= \frac{\frac{1-a}{\tau}\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] - \frac{\kappa}{\tau}\left[(1-\alpha)c\frac{(1-a)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial a}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}\\
&= \frac{\frac{1-a}{\tau}C''(q^*) - \frac{\kappa}{\tau}C'''(q^*)\frac{\partial q^*}{\partial \lambda}}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.
\end{aligned}$$

Because $a < 1$ and $\frac{\partial q^*}{\partial \lambda} > 0$ the cross-partial derivative is positive if $C'''(q^*) = 0$. Otherwise the sign is ambiguous.

Differentiating $\frac{\partial q^*}{\partial \alpha}$ from the proof of Corollary 2 with respect to λ gives

$$\frac{\partial^2 q^*}{\partial \alpha \partial \lambda} = \frac{c\left[\frac{\frac{\partial q^*}{\partial \lambda}\kappa + q^*(1-a)}{\tau}\right]\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right] - c\left(\frac{q^*\kappa}{\tau} + \frac{1}{n}\right)\left[(1-\alpha)c\frac{(1-a)}{\tau} + C'''(q^*)\frac{\partial q^*}{\partial \lambda}\right]}{\left[(1-\alpha)c\frac{\kappa(\lambda,a)}{\tau} + C''(q^*)\right]^2}.$$

Because $a < 1$ and $\frac{\partial q^*}{\partial \lambda} > 0$ the sign of the cross-partial derivative is ambiguous. This completes the proof.

4.A.2 Proofs of the results in Section 4.3.2

Proof of Proposition 2

In this proof, I find the free-entry equilibrium number of health care providers n^* of the game. First, I characterize the equation that defines implicitly n^* as the value of n that solves the equilibrium profit function of parameters $y = (z, K)$ in which $z = (p, \alpha, c, n, \tau, a, \lambda)$. In this characterization, I exploit the monotonicity property of the profit function. Last, I find that profits are decreasing in n^*.

n^* denotes the free-entry equilibrium number of health care providers. I use the assumption of full market coverage. [17]

For a payment scheme $R = (p, \alpha)$, the continuation equilibrium profits for a provider if it enters are the following:

$$\left[p - (1-\alpha)cq^*(n)\right]\frac{1}{n} - C(n) - K, \tag{26}$$

in which $q^*(n)$ is the symmetric equilibrium quality from Corollary 1, $\frac{1}{n}$ is the symmetric demand, and K is the fixed cost of entry. I have assumed that monopoly profits are positive, that is $\frac{p-(1-\alpha)cq^*(1)}{1} - C(q^*(1)) - K > 0$. Then, the first-order derivative of the profits (26) with respect to n becomes

$$-\frac{p-(1-\alpha)cq^*(n)}{n^2} - \left[\frac{(1-\alpha)c}{n} + C'(q^*(n))\right]\frac{\partial q^*(n)}{\partial n}. \tag{27}$$

Using the assumption of $C'(q^*(n) > 0$ and because $\frac{\partial q^*(n)}{\partial n}$ is negative (Corollary 1), (27) is negative. That is, the profits monotonically decrease in n.

Because the profits are positive when n is small enough and they turn negative for large enough n, there is a unique solution n^* for which

$$\pi(n^*, y) \equiv \frac{p-(1-\alpha)cq^*(n^*)}{n^*} - C(q^*(n^*)) - K = 0. \tag{28}$$

in which $y = (z, K)$ and $z = (p, \alpha, c, n, \tau, a, \lambda)$. The equilibrium number of health care providers n^* is defined by $\pi(n^*, y) = 0$. This completes the proof.

Proof of Corollary 5

In this proof I find the comparative statics of the number of providers n^* with respect to the competition parameters K and τ using the implicit function theorem around n^*.

Evaluating (27) at n^* gives $\frac{\partial \pi}{\partial n}$:

$$-\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} - \left[\frac{(1-\alpha)c}{n^*} + C'(q^*(n^*))\right]\frac{\partial q^*(n^*)}{\partial n^*} < 0, \tag{29}$$

---

17 The market is fully covered for every $q \geq 0$ if and only if $\bar{s} + q^* \geq \frac{\tau}{2n^*}$ where the maximum distance to a provider is $\frac{1}{2n*}$. I obtain the market coverage condition by substituting the equilibrium quality q^*, in Proposition 1, and the equilibrium number of providers n^*, characterized in Stage 2, for the market coverage condition $\bar{s} + q^*(n^*) - \frac{\tau}{2n^*} \geq 0$.

because $\frac{\partial q^*(n^*)}{\partial n^*} > 0$. This shows $\frac{\partial \pi}{\partial n} < 0$.

Then I obtain the comparative statics of n^* with respect to entry cost K. Using the implicit function theorem around n^* gives

$$\frac{\partial n^*}{\partial K} = -\frac{\frac{\partial \pi}{\partial K}}{\frac{\partial \pi}{\partial n}} = \frac{-1}{\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} + \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]\frac{\partial q^*(n^*)}{\partial n^*}} < 0,$$

in which (29) is used for $\frac{\partial \pi}{\partial n}$. This gives the first statement of the corollary.

Then, I obtain the comparative statics of n^* with respect to transportation cost τ. Using the implicit function theorem around n^* gives

$$\frac{\partial n^*}{\partial \tau} = -\frac{\frac{\partial \pi}{\partial \tau}}{\frac{\partial \pi}{\partial n}} = \frac{-\left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]\frac{\partial q^*}{\partial \tau}}{\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} + \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]\frac{\partial q^*(n^*)}{\partial n^*}} > 0,$$

in which (29) is used for $\frac{\partial \pi}{\partial n}$ and because $\frac{\partial q^*}{\partial \tau} < 0$. This gives the second result and completes the proof.

Proof of Corollary 6

In this proof I find the comparative statics of the number of providers n^* with respect to the regulation parameters α and p using the implicit function theorem around n^*. Below, I use (29) for $\frac{\partial \pi}{\partial n}$.

I begin by obtaining the comparative statics of n^* with respect to cost reimbursement α. Using the implicit function theorem around n^* gives

$$\frac{\partial n^*}{\partial \alpha} = -\frac{\frac{\partial \pi}{\partial \alpha}}{\frac{\partial \pi}{\partial n}} = \frac{\frac{cq^*(n^*)}{n} - \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]\frac{\partial q^*}{\partial \alpha}}{\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} + \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]}. \tag{30}$$

Then, I obtain the comparative statics of n^* with respect to prospective payment p. Using the implicit function theorem around n^* gives

$$\frac{\partial n^*}{\partial p} = -\frac{\frac{\partial \pi}{\partial p}}{\frac{\partial \pi}{\partial n}} = \frac{\frac{1}{n^*} - \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]\frac{\partial q^*}{\partial p}}{\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} + \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]} \tag{31}$$

The signs of $\frac{\partial n^*}{\partial \alpha}$ and $\frac{\partial n^*}{\partial p}$ are ambiguous because the nominators cannot be signed without imposing further structure to the model.

The first term of the nominator in (30) and (31), $\frac{1}{n}$, gives the direct effect of greater cost sharing and prospective payment to increase entry. The second terms $\left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\right)\right]\frac{\partial q^*}{\partial \alpha}$ and $\left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\right)\right]\frac{\partial q^*}{\partial p}$ are the indirect effects of cost sharing and prospective payment through equilibrium quality. Because $\frac{\partial q^*}{\partial \alpha} > 0$, higher α leads to higher q. Because $\frac{\partial q^*}{\partial p} > 0$, higher p leads to higher q. The signs of $\frac{\partial n^*}{\partial \alpha}$ and $\frac{\partial n^*}{\partial p}$ depends on which of these effects are bigger.

Because $\frac{\partial q^*}{\partial \alpha} = \frac{c\left[\frac{q^*}{\tau}+\frac{1}{n}\right]}{(1-\alpha)c\frac{\kappa}{\tau}+C''(q^*)}$, the effect of $\frac{\partial q^*}{\partial \alpha}$ is small if the cost of providing higher quality is very convex (C'' is large). Then, the direct effect dominates, and entry increases with the cost sharing.

Because $\frac{\partial q^*}{\partial p} = \frac{\frac{\kappa}{\tau}}{(1-\alpha)c\frac{\kappa}{\tau}+C''(q^*)}$, the effect of $\frac{\partial q^*}{\partial p}$ is small if the cost of providing higher quality is very convex (C'' is large). Then, the direct effect dominates, and entry increases with the prospective payment.

Thus, the two effects in (30) and (31) are positive if the indirect cost effect, caused by a higher equilibriun quality, is smaller than the profitability effect.

Proof of Corollary 7

In this proof I find the comparative statics of the number of providers n^* with respect to the patient misperception parameters a and λ. These are obtained by using the implicit function theorem around n^*.

I begin by obtaining the comparative statics of n^* with respect to patients' responsiveness on quality a. Using the implicit function theorem around n^* gives

$$\frac{\partial n^*}{\partial a} = -\frac{\frac{\partial \pi}{\partial a}}{\frac{\partial \pi}{\partial n}} = \frac{-\left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\right)\right]\frac{\partial q^*}{\partial a}}{\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} + \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]} < 0,$$

in which (29) is used for $\frac{\partial \pi}{\partial n}$ and because $\frac{\partial q^*}{\partial a} > 0$ (Corollary 3).

Next I obtain the comparative statics of n^* with respect to the share of patients with accurate perceptions λ and consider $a > 1$ first. Using the implicit function theorem around n^* gives

$$\frac{\partial n^*}{\partial \lambda} = -\frac{\frac{\partial \pi}{\partial \tau}}{\frac{\partial \pi}{\partial n}} = \frac{-\left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]\frac{\partial q^*}{\partial \lambda}}{\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} + \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]} > 0$$

in which (29) is used for $\frac{\partial \pi}{\partial n}$ and because $\frac{\partial q^*}{\partial \lambda} < 0$ (Corollary 3).

Next I obtain the comparative statics of n^* with respect to the share of patients with accurate perceptions λ. For $a < 1$

$$\frac{\partial n^*}{\partial \lambda} = -\frac{\frac{\partial \pi}{\partial \tau}}{\frac{\partial \pi}{\partial n}} = \frac{-\left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\right)\right]\frac{\partial q^*}{\partial \lambda}}{\frac{p-(1-\alpha)cq^*(n^*)}{n^{*2}} + \left[\frac{(1-\alpha)c}{n^*} + C'\left(q^*\left(n^*\right)\right)\right]} < 0$$

in which (29) is used for $\frac{\partial \pi}{\partial n}$ and because $\frac{\partial q^*}{\partial \lambda} > 0$ (Corollary 3).

4.A.3 Proofs of the results in Section 4.4.1

Derivation of regulator's objective function

Here I derive the regulator's objective function for the analysis in Section 4.4. I consider a utilitarian social welfare function in which the social planner gives equal weights to patients and providers. The social welfare function is the sum of the patient surplus (consumer surplus, CS) and the producer surplus (PS) minus the costs of the regulated payment scheme (RC):

$$W(Q,n) = CS(Q,n) + PS(Q,n) - RC(Q,n), \tag{32}$$

in which Q denotes the vector of qualities and n is the number of providers. I consider each part separately and begin with the patient surplus.

The patient surplus consists of the utility of receiving the treatment subtracted by the disutility of the transportation costs

$$CS(Q,n) = \bar{s} + q - \tau \left[2n \int_0^{1/2n} x dx \right], \tag{33}$$

where $2\tau n \int_0^{1/2n} x dx$ are the total transportation costs. (33) becomes

$$CS(Q,n) = \bar{s} + q - \frac{\tau}{4n}. \tag{34}$$

I consider provider surplus next. Because there are n providers in the market, the provider surplus that is the sum of provider profits is

$$PS(Q,n) = n\left[\pi(Q,n) - K\right]. \tag{35}$$

Because I consider a symmetric equilibrium, the demand is $1/n$, and I substitute this for (5) so (36) becomes

$$PS(Q,n) = n\left\{ \left[p - (1-\alpha)cq\right]\frac{1}{n} - C(q) - K \right\}. \tag{36}$$

I simplify (36) to become

$$PS(Q,n) = p - (1-\alpha)cq - nC(q) - nK. \tag{37}$$

Last, I derive the regulatory costs following Bardey et al. (2012). Again, I use the symmetry to simplify demand. Because a lump sum is levied to finance the regulatory costs and the regulator does not reimburse the fixed quality production costs of quality, $C(q)$, or the entry cost, K, the regulatory costs are as follows:

$$\begin{aligned} RC(Q,n) &= n\left\{ \left[p + (c - (1-\alpha)c)q\right]\frac{1}{n} \right\} \\ &= p + \left[c - (1-\alpha)c\right] q. \end{aligned} \tag{38}$$

Therefore, using (34), (37), and (38) for (32) for any symmetric equilibrium, the social welfare is

$$W(Q,n) = \bar{s} + q - \frac{\tau}{4n} - nC(q) - nK - cq. \tag{39}$$

(39) is (12).[18] This completes the derivation.

18 Using (34), (37), and (38) for (32) social welfare is $W(Q_s, n_s) = \{\bar{s} + q - \frac{\tau}{4n}\} + \{p - (1-\alpha)cq - nC(q) - nK\} - \{p + (c - \hat{c})q\}$, which becomes (39).

Proof of Proposition 3

In this proof I characterize the first best quality and number of providers, q^{FB} and n^{FB}. Differentiating (12) with respect to q

$$1 - c - nC'(q). \tag{40}$$

Differentiating (12) with respect to n gives

$$\frac{4\tau}{16n^2} - C(q) - K. \tag{41}$$

Setting these first-order derivatives to zero gives (13) and (14). The first-order conditions (13) and (14) implicitly define q^{FB} and n^{FB} as functions of (c, τ, K). This completes the proof.

Proof of Corollary 8

In this proof I find the comparative statics of q^{FB} and n^{FB} with respect to c, τ, and K. Suppose (13) and (14) are satisfied at the point $(q^{FB}, n^{FB}; c, \tau, K)$. The partial derivatives of O^1 and O^2 are continuous. I use the implicit function theorem and Cramer's rule to find the partial derivatives of q^{FB} and n^{FB} with respect to c, τ, and K.

I take the total differential of (13) and (14) with respect to endogenous and exogenous variables and obtain

$$-nC''(q)\mathrm{d}q - C'(q)\mathrm{d}n - 1\mathrm{d}c = 0 \tag{42}$$

$$-C'(q)\mathrm{d}q - \frac{\tau}{2n^3}\mathrm{d}n + \frac{1}{4n^2}\mathrm{d}\tau - 1\mathrm{d}K = 0 \tag{43}$$

In matrix form, (42) and (43) are

$$\begin{bmatrix} -nC''(q) & -C'(q) & -1 & 0 & 0 \\ -C'(q) & -\frac{\tau}{2n^3} & 0 & \frac{1}{4n^2} & -1 \end{bmatrix} \begin{bmatrix} \mathrm{d}q \\ \mathrm{d}n \\ \mathrm{d}c \\ \mathrm{d}\tau \\ \mathrm{d}K \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The Jacobian matrix is

$$\left|J\right| \equiv \begin{vmatrix} -nC''(q) & -C(q) \\ -C'(q) & -\frac{\tau}{2n^3} \end{vmatrix}.$$

The Hessian matrix of the objective function has to be negative definite for (13) and (14) to characterize a local maximum in the first best. Thus, I make an additional assumption that the determinant of matrix J is positive. This would follow, for example, if C is convex enough, that is C'' is large enough. Assuming, $\left|J\right| > 0$, I can apply the implicit function theorem and consider the effect of each exogenous parameter c, τ, and K in turn by using the Cramer's rule.

The partial derivative of q^{FB} with respect to c is

$$\frac{\partial q^{FB}}{\partial c} = \frac{\begin{vmatrix} 1 & -C'(q) \\ 0 & -\frac{\tau}{2n^3} \end{vmatrix}}{\left|J\right|} = \frac{-\dfrac{\tau}{2n^3}}{\left|J\right|} < 0.$$

The partial derivative of n^{FB} with respect to c is

$$\frac{\partial n^{FB}}{\partial c} = \frac{\begin{vmatrix} -nC''(q) & 1 \\ -C'(q) & 0 \end{vmatrix}}{\left|J\right|} = \frac{C'(q)}{\left|J\right|} > 0,$$

because $\left|J\right| > 0$ and $C'(q) > 0$. The partial derivative of q^{FB} with respect to τ is

$$\frac{\partial q^{FB}}{\partial \tau} = \frac{\begin{vmatrix} 0 & -C'(q) \\ -\frac{1}{4n^2} & -\frac{\tau}{2n^3} \end{vmatrix}}{\left|J\right|} = \frac{-\frac{C'(q)}{4n^2}}{\left|J\right|} < 0.$$

The partial derivative of n^{FB} with respect to τ is

$$\frac{\partial n^{FB}}{\partial \tau} = \frac{\begin{vmatrix} -nC'(q) & 0 \\ -C'(q) & -\frac{1}{4n^2} \end{vmatrix}}{\left|J\right|} = \frac{\frac{C'(q)}{4n}}{\left|J\right|} > 0,$$

because $\left|J\right| > 0$ and $C'(q) > 0$. The partial derivative of q^{FB} with respect to K is

$$\frac{\partial q^{FB}}{\partial K} = \frac{\begin{vmatrix} 0 & -C(q) \\ 1 & -\frac{\tau}{2n^3} \end{vmatrix}}{\left|J\right|} = \frac{C(q)}{\left|J\right|} > 0.$$

The partial derivative of n^{FB} with respect to K is

$$\frac{\partial n^{FB}}{\partial K} = \frac{\begin{vmatrix} -nC''(q) & 0 \\ -C'(q) & 1 \end{vmatrix}}{\left|J\right|} = \frac{-nC''(q)}{\left|J\right|} < 0,$$

because $\left|J\right| > 0$. This completes the proof.

Proof of Corollary 9

I have omitted the proof because the statement in Corollary 9 follows immediately from what is stated in the main text.

4.A.4 Proofs of the results in Section 4.4.2

Proof of Proposition 4

In this proof I find the prospective payment $\widehat{p}$ that allows the regulator to implement the first best quality by granting licenses for n^{FB} providers. Last I show that the equilibrium q is unique and that $q = q^{FB}$.[19]

I begin the proof by considering some fixed $\left(q^{FB}, n^{FB}\right)$. Recall that $\left(q^{FB}, n^{FB}\right)$ is obtained implicitly as a solution to equations (13) and (14). I evaluate (13) and (14) at $\left(q^{FB}, n^{FB}\right)$ and re-arrange the resulting equations and obtain

$$\frac{1}{n^{FB}} = \frac{c}{n^{FB}} + C'(q^{FB}) \tag{44}$$

and

$$\frac{\tau}{4(n^{FB})^2} = C(q^{FB}) + K. \tag{45}$$

Using the assumption that the regulator uses the prospective payment as its payment instrument only, I can simplify (7), which determines the equilibrium quality as follows:

$$\tilde{G}(q;z) \equiv (p - cq)\frac{\kappa}{\tau} - \frac{c}{n} - C'(q) = 0 \tag{46}$$

Re-arranging (46) gives:

$$[p - cq]\,\frac{\kappa}{\tau} = \frac{c}{n} + C'(q). \tag{47}$$

The regulator can use (44) and (45) to calculate what level of equality, q^{FB}, and entry, n^{FB}, achieves the first best. Because entry is contractible, the regulator can set $n = n^{FB}$. Now, solving (47) for p gives

$$p = cq + \frac{1}{n}\frac{\tau}{\kappa}. \tag{48}$$

Imposing (q^{FB}, n^{FB}) to (48) gives the prospective payment $\widehat{p}$ which the regulator can use to implement the first-best, given that it can directly set $n = n^{FB}$

$$\widehat{p} = cq^{FB} + \frac{1}{n^{FB}}\frac{\tau}{\kappa}. \tag{49}$$

This gives (15) in Proposition 4.

To see that $\widehat{p}$ indeed results in the first best and that q^{FB} and n^{FB} are chosen by the providers in equilibrium, I substitute the prospective payment $\widehat{p}$ from (49) for (47)

$$\left[cq^{FB} + \frac{1}{n^{FB}}\frac{\tau}{\kappa} - cq\right]\frac{\kappa}{\tau} = \frac{c}{n^{FB}} + C'(q). \tag{50}$$

[19] I have omitted the arguments λ and a from κ to simplify the notation.

Then I re-arrange (50) and take common terms to define the following function, $A(q)$:

$$A(q) = \left[q^{FB} - q\right] \frac{\kappa}{\tau} c + \frac{1}{n^{FB}} - \frac{c}{n^{FB}} - C'(q). \tag{51}$$

The level of quality that solves $A(q) = 0$ is the equilibrium quality chosen by the providers (see (7)). Evaluating (51) at q^{FB} shows that

$$A(q^{FB}) \equiv \frac{1}{n^{FB}} - \frac{c}{n^{FB}} - C'(q^{FB}) = 0. \tag{52}$$

Notice also that

$$A(0) = q^{FB} \frac{\kappa}{\tau} c + (1 - c) \frac{1}{n^{FB}} > 0, \tag{53}$$

because $C'(0) = 0$ by assumption. In addition, $\lim_{q \to \infty} A(q) < 0$, because $\lim_{q \to \infty} C'(q) = -\infty$.

Then, using the result of $A(0) > 0$ given by (53) and $\lim_{q \to \infty} C'(q) = -\infty$, there is a unique q that solves $A(q) = 0$, and this unique q is both equal to q^* and q^{FB}. This completes the proof.

Proof of Corollary 10

In this proof, I find the comparative statics of $\widehat{p}$ with respect to K, τ and c.[20]

Total differentiation of (15) with respect to K and using the results from Corollary 8 gives

$$\frac{\partial \widehat{p}}{\partial K} = c \underbrace{\frac{\partial q^{FB}}{\partial K}}_{>0} - \frac{\tau}{(\kappa n^{FB})^2} \underbrace{\frac{\partial n^{FB}}{\partial K}}_{<0} > 0.$$

Total differentiation of (15) with respect to τ and using the results from Corollary 8 gives

$$\frac{\partial \widehat{p}}{\partial \tau} = c \underbrace{\frac{\partial q^{FB}}{\partial \tau}}_{<0} - \frac{\tau}{(\kappa n^{FB})^2} \underbrace{\frac{\partial n^{FB}}{\partial \tau}}_{>0} + \frac{1}{n\kappa}.$$

Total differentiation of (15) with respect to c and using the results from Corollary 8 gives

$$\frac{\partial \widehat{p}}{\partial c} = q^{FB} + \underbrace{\frac{\partial q^{FB}}{\partial c}}_{<0} + \frac{\tau}{(\kappa n^{FB})^2} \underbrace{\frac{\partial n^{FB}}{\partial c}}_{>0}.$$

These give the results stated in the corollary and thus complete the proof.

20 I have omitted the arguments λ and a from κ to simplify the notation.

Proof of Corollary 11

In this proof, I find the comparative statics of $\widehat{p}$ with respect to misperception parameters a and α.

Applying $\frac{\partial \kappa(\lambda,a)}{\partial a} = 1 - \lambda$ and Corollary 9, total differentiation of (15) with respect to a gives

$$\frac{\partial \widehat{p}}{\partial a} = -\frac{(1-\lambda)\tau n^{FB}}{[\kappa(\lambda,a)n^{FB}]^2} < 0. \tag{54}$$

Applying $\frac{\partial \kappa(\lambda,a)}{\partial \lambda} = 1 - a$ and Corollary 9, total differentiation of (15) with respect to λ gives

$$\frac{\partial \widehat{p}}{\partial \lambda} = -\frac{(1-a)\tau n^{FB}}{[\kappa(\lambda,a)n^{FB}]^2}. \tag{55}$$

If some patients overreact to quality $(a > 1)$, (55) is positive. If some patients underreact to quality $(a < 1)$, (55) is negative. This completes the proof.

4.A.5 Proofs of the results in Section 4.4.3

Proof of Proposition 5

In this proof I find the prospective payment p^A and cost reimbursement α^A that allows the regulator to implement the first best when Regime A prevails and when there is an interior solution to the regulator's maximization problem. The proof consists of three parts. After describing the regulator's maximization problem I find the first-order derivatives with respect to p and α and give the first order conditions. Second, I check that p^A and α^A induce the providers to choose the first best. Last, I check that the solution (p^A, α^A) is also an interior solution assuming that such a solution exists.[21]

The regulator chooses a payment scheme $R = (p, \alpha)$ that maximizes social welfare. Given that the continuation equilibrium number of the providers is n^* and qualities q^*, the parameters p and α are solutions to the following problem:

$$\max_{R=(p,\alpha)} \bar{s} + q^*\left(n^*(R), R\right) - \frac{\tau}{4n^*(R)} - n^*(R)C\left(q^*\left(n^*(R),R\right)\right) - Kn^*(R) - cq^*\left(n^*(R),R\right) \tag{56}$$

subject to $\alpha \geq 0$ and $\alpha \leq 1$. The Lagrangian function for the optimization problem (74) is

$$\begin{aligned} L = \bar{s} + q^*\left(n^*(R), R\right) - \frac{\tau}{4n^*(R)} - n^*(R)C\left(q^*\left(n^*(R),R\right)\right) - Kn^*(R) - cq^*\left(n^*(R),R\right) \\ - \nu\left[(1-\alpha) - 1\right] + \mu\left[1-\alpha\right] \end{aligned} \tag{57}$$

[21] I have omitted the arguments λ and a from κ to simplify the notation.

Multipliers $\mu > 0$ and $\nu = 0$ correspond to full cost reimbursement. In this case $\alpha = 1$. Multipliers $\nu > 0$ and $\mu = 0$ imply pure prospective payment. In this case $\alpha = 0$.[22]

I differentiate the Lagrangian given in (57). The first-order conditions with respect to p and α are

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R), R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}p} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R), R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}p} = 0 \tag{58}$$

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R), R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R), R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}\alpha} + \nu - \mu = 0, \tag{59}$$

where $\frac{\mathrm{d}q^*}{\mathrm{d}p} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial p} + \frac{\partial q^*}{\partial p}$ and $\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial \alpha} + \frac{\partial q^*}{\partial \alpha}$.

In this proof I focus on Regime A when neither of the constraints bind, $\nu = \mu = 0$. In this regime, the regulator uses a mixed payment instrument that consists of a prospective payment and a partial cost reimbursement so that $\alpha \in (0, 1)$. Thus, (76) and (77) become

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R), R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}p} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R), R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}p} = 0 \tag{60}$$

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R), R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R), R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}\alpha} = 0, \tag{61}$$

where $\frac{\mathrm{d}q^*}{\mathrm{d}p} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial p} + \frac{\partial q^*}{\partial p}$ and $\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial \alpha} + \frac{\partial q^*}{\partial \alpha}$.

Next, I find the equilibrium prospective payment p^A and cost reimbursement α^A that yield q^{FB} and n^{FB} in Regime A. I take q^{FB} and n^{FB}, a solution to (13) and (14). Then I substitute q^{FB} and n^{FB} for q and n in the equations that determine the equilibrium quality and the equilibrium number of providers (equations (7) and (10)). After substituting q^{FB} and n^{FB} for (7) and (10), there will be two equations with two unknowns (α, p). These are

$$\left[p - (1-\alpha)cq^{FB}\right]\frac{\kappa}{\tau} - \frac{(1-\alpha)c}{n^{FB}} - C'\left(q^{FB}\right) = 0 \tag{62}$$

and

$$\left[\frac{(1-\alpha)c}{n^{FB}} + C'(q^{FB})\right]\frac{\tau}{\kappa n^{FB}} - C\left(q^{FB}\right) - K = 0. \tag{63}$$

When solved, I get the following expression for p:

$$p = (1-\alpha)cq^{FB} + \left[\frac{(1-\alpha)c}{n^{FB}} + C'(q^{FB})\right]\frac{\tau}{\kappa}. \tag{64}$$

22 For L I also use (implicitly) the earlier assumption of non-negative prospective payment.

Then, I obtain the following result from (63):

$$(1-\alpha)c = \left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB} \tag{65}$$

I use the result given by (65) when finding the explicit solution for p^A. Next, I solve α from (65)

$$\alpha^A \equiv 1 - \frac{\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}}{c}, \tag{66}$$

in which α^A is the equilibrium cost reimbursement in Regime A.

Last, I find the explicit solution for p^A. I begin by re-arranging (64) as follows:

$$p = (1-\alpha)c\left[q^{FB}+\frac{\tau}{\kappa n^{FB}}\right]+C'(q^{FB})\frac{\tau}{\kappa} \tag{67}$$

Then, I substitute $(1-\alpha)c$ from (65) for (64)[23]

$$p = \left\{\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}\right\}\left[q^{FB}+\frac{\tau}{\kappa n^{FB}}\right]+C'(q^{FB})\frac{\tau}{\kappa}. \tag{68}$$

Re-arranging (62) gives

$$p^A \equiv \left\{\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}\right\}q^{FB}+\left[C\left(q^{FB}\right)+K\right]n^{FB}. \tag{69}$$

Equations (66) and (69) give the first and second equations in Proposition 3.

To check that p^A and α^A induce the providers to choose the first best, I substitute them for (7) and (10). Then I obtain the following two equations, in which the unknowns are (q, n)

$$\begin{aligned}&\left\{\left\{\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}\right\}q^{FB}+\left[C\left(q^{FB}\right)+K\right]n^{FB}\right.\\&\left.-\left\{\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}\right\}q\right\}\frac{\kappa}{\tau}\\&-\frac{\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}}{n}-C'\left(q\right)=0\end{aligned} \tag{70}$$

and

$$\begin{aligned}&\left\{\left[\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}\right]q^{FB}+\left[C\left(q^{FB}\right)+K\right]n^{FB}\right.\\&\left.-\left[\left[C\left(q^{FB}\right)+K\right]\frac{\kappa}{\tau}(n^{FB})^2 - C'(q^{FB})n^{FB}\right]q\right\}\frac{1}{n}-C\left(q\left(n\right)\right)-K=0.\end{aligned} \tag{71}$$

These two equations show how the equilibrium (q, n) are determined when the regulator sets (p^A, α^A), given by (70) and (71). In particular if the regulator sets

[23] Instead of substituting α using (66), it is convenient to substitute $(1-\alpha)c$ from (65) instead. Using α from (66) yields the same solution for (69).

(p^A, α^A), I obtain $q^* = q^{FB}$ and $n^* = n^{FB}$. These values are such that (61) and (60) are equal to zero. This can be seen by comparing the terms in square brackets of (61) and (60) to equations (14) and (13) in the main text. Thus,

$$\frac{\mathrm{d}W}{\mathrm{d}p} = 0$$
$$\frac{\mathrm{d}W}{\mathrm{d}\alpha} = 0,$$

which means that the solution is also an interior solution to the welfare maximization problem, assuming that such a solution exists. This completes the proof.

Proof of Corollary 12

In this proof, I find the partial derivatives of p^A and α^A with respect to a, λ, K, τ, and c.

I begin by finding the partial derivatives with respect to misperception parameters (a and λ). Because q^{FB} and n^{FB} are independent of a and λ (Corollary 9), total differentiation of p^A with respect to a gives

$$\frac{\partial p^A}{\partial a} = \left\{ \left[C(q^{FB}) + K \right] \frac{1-\lambda}{\tau} \left(n^{FB} \right)^2 \right\} q^{FB} > 0.$$

Similarly, using Corollary 9, total differentiation of p^A with respect to λ gives

$$\frac{\partial p^A}{\partial \lambda} = \left\{ \left[C(q^{FB}) + K \right] \frac{1-a}{\tau} \left(n^{FB} \right)^2 \right\} q^{FB}. \tag{72}$$

If some patients overreact to quality ($a > 1$), (72) is negative. If some patients underreact to quality ($a < 1$), (72) is positive.

Because q^{FB} and n^{FB} are independent of a and λ (Corollary 9), total differentiation the partial derivative of α^A with respect to a gives

$$\frac{\partial \alpha^A}{\partial a} = -\left\{ \left\{ \left[C(q^{FB}) + K \right] \frac{1-\lambda}{\tau} \left(n^{FB} \right)^2 \right\} q^{FB} \right\} \frac{1}{c} < 0.$$

Using Corollary 9, total differentiation the partial derivative of α^A with respect to λ gives

$$\frac{\partial \alpha^A}{\partial \lambda} = -\left\{ \left\{ \left[C(q^{FB}) + K \right] \frac{1-\lambda}{\tau} \left(n^{FB} \right)^2 \right\} q^{FB} \right\} \frac{1}{c}. \tag{73}$$

If some patients overreact to quality ($a > 1$), (73) is positive. If some patients underreact to quality ($a < 1$), (73) is negative.

Total differentiation of p^A from Proposition 5 with respect to K and using the results from Corollary 8 gives

$$\begin{aligned}\frac{\partial p^A}{\partial K} =& \Big\{ \Big[C'(q^{FB})\frac{\partial q^{FB}}{\partial K}+1\Big]\frac{\kappa}{\tau}(n^{FB})^2+\Big[C(q^{FB})+K\Big]\frac{\kappa}{\tau}2n^{FB}\frac{\partial n^{FB}}{\partial K} \\ &-C''(q^{FB})n^{FB}\frac{\partial q^{FB}}{\partial K}-C'(q^{FB})\frac{\partial n^{FB}}{\partial K}\Big\}q^{FB}+\Big\{\Big[C(q^{FB})+K\Big]\frac{\kappa}{\tau}(n^{FB})^2-C'(q^{FB})n^{FB}\Big\}\frac{\partial q^{FB}}{\partial K} \\ &+\Big\{C'(q^{FB})\frac{\partial q^{FB}}{\partial K}+1\Big\}n^{FB}+\Big[C(q^{FB})+K\Big]\frac{\partial n^{FB}}{\partial K}.\end{aligned}$$

Total differentiation of p^A from Proposition 5 with respect to τ and using the results from Corollary 8 gives

$$\begin{aligned}\frac{\partial p^A}{\partial \tau} =& \Big\{\frac{\kappa(n^{FB})^2\left[C(q^{FB})+K\right]}{\tau}-C'(q^{FB})n^{FB}\Big\}\frac{\partial q^{FB}}{\partial \tau} \\ &+\Big\{\frac{\tau\Big(2\kappa n^{FB}\frac{\partial n^{FB}}{\partial \tau}\left[C(q^{FB})+K\right]+\kappa(n^{FB})^2C'(q^{FB})\frac{\partial q^{FB}}{\partial \tau}\Big)-\kappa(n^{FB})^2\left[C(q^{FB})+K\right]}{\tau^2} \\ &-C''(q^{FB})\frac{\partial q^{FB}}{\partial \tau}n^{FB}-C'(q^{FB})\frac{\partial n^{FB}}{\partial \tau}\Big\}q^{FB}+C'(q^{FB})\frac{\partial q^{FB}}{\partial \tau}n^{FB} \\ &+\Big[C(q^{FB})+K\Big]\frac{\partial n^{FB}}{\partial \tau}.\end{aligned}$$

Total differentiation of p^A from Proposition 5 with respect to c and using the results from Corollary 8 gives

$$\begin{aligned}\frac{\partial p^A}{\partial c} =& \Big\{\frac{\kappa(n^{FB})^2\left[C(q^{FB})+K\right]}{\tau}-C'(q^{FB})n^{FB}\Big\}\frac{\partial q^{FB}}{\partial c} \\ &+\Big\{\frac{2\kappa n^{FB}\frac{\partial n^{FB}}{\partial c}\left[C(q^{FB})+K\right]+\kappa(n^{FB})^2C'(q^{FB})\frac{\partial q^{FB}}{\partial c}}{\tau} \\ &-C''(q^{FB})\frac{\partial q^{FB}}{\partial c}n^{FB}-C'(q^{FB})\frac{\partial n^{FB}}{\partial c}\Big\}q^{FB}+C'(q^{FB})\frac{\partial q^{FB}}{\partial c}n^{FB}+\Big[C(q^{FB})+K\Big]\frac{\partial n^{FB}}{\partial c}.\end{aligned}$$

Total differentiation of α^A from Proposition 5 with respect to K and using the results from Corollary 8 gives

$$\begin{aligned}\frac{\partial \alpha^A}{\partial K} =& -\frac{1}{c}\Big\{\Big[C'(q^{FB})\frac{\partial q^{FB}}{\partial K}+1\Big]\frac{\kappa}{\tau}(n^{FB})^2+\Big[C(q^{FB})+K\Big]\frac{\kappa}{\tau}2n^{FB}\frac{\partial n^{FB}}{\partial K} \\ &-C''(q^{FB})\frac{\partial q^{FB}}{\partial K}n^{FB}-C'(q^{FB})\frac{\partial n^{FB}}{\partial K}\Big\}.\end{aligned}$$

Total differentiation of α^A from Proposition 5 with respect to τ and using the results from Corollary 8 gives

$$\begin{aligned}\frac{\partial \alpha^A}{\partial \tau} =& -\frac{1}{c}\Big\{\frac{\tau(2\kappa n^{FB}\frac{\partial n^{FB}}{\partial \tau}\left[C(q^{FB})+K\right]+\kappa(n^{FB})^2C'(q^{FB})\frac{\partial q^{FB}}{\partial \tau})-\kappa(n^{FB})^2\left[C(q^{FB})+K\right]}{\tau^2} \\ &-C''(q)\frac{\partial q^{FB}}{\partial \tau}n^{FB}-C'(q^{FB})\frac{\partial n^{FB}}{\partial \tau}\Big\}.\end{aligned}$$

Total differentiation of α^A from Proposition 5 with respect to c and using the results from Corollary 8 gives

$$\begin{aligned}\frac{\partial \alpha^A}{\partial c} =& \frac{1}{c^2}\left\{\frac{\kappa(n^{FB})^2\left[C(q^{FB})+K\right]}{\tau} - C'(q^{FB})n^{FB}\right\}\\ &- \frac{1}{c}\left\{\frac{2\kappa n^{FB}\frac{\partial n^{FB}}{\partial c}\left[C(q^{FB})+K\right]+\kappa(n^{FB})^2C'(q^{FB})\frac{\partial q^{FB}}{\partial c}}{\tau} - C''(q^{FB})\frac{\partial q^{FB}}{\partial c}n^{FB}\right.\\ &\left.- C'(q^{FB})\frac{\partial n^{FB}}{\partial c}\right\}.\end{aligned}$$

The partial derivatives show that the effects of exogenous parameters are ambiguous.

Proof of Proposition 6

In this proof I show that if Regime B or Regime C prevails, the first best cannot be implemented. I begin by describing the maximization problem of the regulator first. Then I consider each of the regimes separately and use proof by contradiction that the claim in the proposition is true.[24]

The regulator chooses a payment scheme $R = (p, \alpha)$ that maximizes social welfare. Given the continuation equilibrium number of providers n^* and qualities q^*, the parameters p and α are solutions to the following problem:

$$\begin{aligned}\max_{R=(p,\alpha)} \bar{s} + q^*\left(n^*(R), R\right) - \frac{\tau}{4n^*(R)} - n^*(R)C\left(q^*\left(n^*(R), R\right)\right)\\ -Kn^*(R) - cq^*\left(n^*(R), R\right)\end{aligned} \tag{74}$$

subject to $\alpha \geq 0$ and $\alpha \leq 1$. The Lagrangian function for the optimization problem (74) is

$$\begin{aligned}L = \bar{s} + q^*\left(n^*(R), R\right) - \frac{\tau}{4n^*(R)} - n^*(R)C\left(q^*\left(n^*(R), R\right)\right) - Kn^*(R)\\ - cq^*\left(n^*(R), R\right) - \nu\left[(1-\alpha)-1\right] + \mu\left[1-\alpha\right]\end{aligned} \tag{75}$$

In Regime B, $\alpha < 1$ constraint binds, that is $\nu > 0$ and $\mu = 0$. Multipliers $\nu > 0$ and $\mu = 0$ imply pure prospective payment. In Regime C, $\alpha = 1$ constraint binds, that is $\nu = 0$ and $\mu > 0$. Multipliers $\nu = 0$ and $\mu > 0$ correspond to full cost reimbursement.[25]

I differentiate the Lagrangian in (75) to obtain the following first-order con-

24 I have omitted the arguments λ and a from κ to simplify the notation.

25 I also use (implicitly) the earlier assumption of non-negative prospective payment for L.

ditions with respect to p and α

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R),R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}p} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R),R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}p} = 0 \tag{76}$$

$$n^*(R)\left[\frac{1-c}{n^*(R)} - C'\left(q^*\left(n^*(R),R\right)\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} + \left[\frac{\tau}{4n^{*2}} - C\left(q^*\left(n^*(R),R\right)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}\alpha} + \nu - \mu = 0, \tag{77}$$

where $\frac{\mathrm{d}q^*}{\mathrm{d}p} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial p} + \frac{\partial q^*}{\partial p}$ and $\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial \alpha} + \frac{\partial q^*}{\partial \alpha}$.

First, I prove the result regarding Regime B by using proof by contradiction. Assume, for a contradiction, that the upper constraint on α binds, that is $\alpha = 0$ and that the first-best is achieved, that is $q^* = q^{FB}$ and $n^* = n^{FB}$. The first-order derivatives (76) and (77) become

$$n^{FB}\left[\frac{1-c}{n^{FB}} - C'\left(q^{FB}(n^{FB}(R))\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}p} + \left[\frac{\tau}{(4n^{FB})^2} - C\left(q^{FB}(n(R),R)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}p} = 0 \tag{78}$$

$$n^{FB}\left[\frac{1-c}{n^{FB}} - C'\left(q^{FB}(n^{FB}(R))\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} + \left[\frac{\tau}{4n^{FB2}} - C\left(q^{FB}(n(R),R)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}\alpha} + \nu = 0 \tag{79}$$

where $\frac{\mathrm{d}q^*}{\mathrm{d}p} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial p} + \frac{\partial q^*}{\partial p}$ and $\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial \alpha} + \frac{\partial q^*}{\partial \alpha}$. Evaluating (13) and (14) in (q^{FB}, n^{FB}) and using Proposition 3, the terms in square brackets in (78) and (79) are zero. As shown, imposing $q^* = q^{FB}$ and $n^* = n^{FB}$ to these conditions leads to a contradiction because $\nu > 0$ if the upper constraint on α binds. Therefore, it cannot hold that when $\alpha = 0$, $q^* = q^{FB}$ and $n^* = n^{FB}$ in equilibrium. This gives the first statement in Proposition 6.

Second, I prove the result regarding Regime C using proof by contradiction. Assume, for a contradiction, that the upper constraint on α binds, that is $\alpha = 1$ and that the first-best is achieved, that is $q^* = q^{FB}$ and $n^* = n^{FB}$. The first-order derivatives (76) and (77) become

$$n^{FB}\left[\frac{1-c}{n^{FB}} - C'\left(q^{FB}(n^{FB}(R))\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}p} + \left[\frac{\tau}{(4n^{FB})^2} - C\left(q^{FB}(n(R),R)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}p} = 0 \tag{80}$$

$$n^{FB}\left[\frac{1-c}{n^{FB}} - C'\left(q^{FB}(n^{FB}(R))\right)\right]\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} + \left[\frac{\tau}{4n^{FB2}} - C\left(q^{FB}(n(R),R)\right) - K\right]\frac{\mathrm{d}n^*}{\mathrm{d}\alpha} - \mu = 0 \tag{81}$$

where $\frac{\mathrm{d}q^*}{\mathrm{d}p} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial p} + \frac{\partial q^*}{\partial p}$ and $\frac{\mathrm{d}q^*}{\mathrm{d}\alpha} = \frac{\partial q^*}{\partial n}\frac{\partial n^*}{\partial \alpha} + \frac{\partial q^*}{\partial \alpha}$. Evaluating (13) and (14) in (q^{FB}, n^{FB}) and using Proposition 3, the terms in square brackets in (80) and (81) are zero. As shown, imposing $q^* = q^{FB}$ and $n^* = n^{FB}$ to these conditions leads to a contradiction because $\mu > 0$ if the the upper constraint on α binds. Therefore, it cannot hold that $\alpha = 1$ and $q^* = q^{FB}$ and $n^* = n^{FB}$ in equilbrium. This gives the second statement in Proposition 6 and completes the proof.

4.A.6 Example

Here I provide an example of the model using a more specific form on the per-unit production cost of quality C. This example illustrates why analyzing the general model in my main text is more useful than analyzing the explicit solutions for equilibria.

Let the fixed costs be quadratic $C(q) = \frac{\gamma}{2}q^2$, with $\gamma > 0$ being the weight on the fixed production costs of quality. Then in subgame (R, n) the explicit solution of the symmetric subgame equilibrium quality q^* (Proposition 1) is

$$q^* = \frac{p\kappa\left(\lambda, a\right) - \tau\frac{\widehat{c}}{n}}{\widehat{c}\kappa\left(\lambda, a\right) + \tau\gamma}. \tag{82}$$

Then the symmetric equilibrium of the two-stage entry-quality game is represented by a vector $Q^*\left(R\right) = \left(q^*\left(R\right), ..., q^*\left(R\right)\right)$, in which $q^*\left(R\right)$ is given by (82), and

$$n^*\left(R\right) = \frac{2p\gamma\tau\widehat{c}\kappa + p\gamma^2\tau^2 + \sqrt{\Omega\left(\widehat{c}, K, \kappa, \tau, p\right),}}{\gamma p^2\kappa^2 + 2K\left(\widehat{c}\kappa + \gamma\tau\right)^2}, \tag{83}$$

in which $\Omega\left(\widehat{c}, K, \kappa, \tau, p\right) = \tau\left(2K\widehat{c}^2 + \gamma p^2\right)\left(\widehat{c}\kappa + \gamma\tau\right)^2\left(2\widehat{c}\kappa + \gamma\tau\right)$, from which a and λ have been omitted from $\kappa\left(a, \lambda\right)$.

The expressions above illustrate that even with a simple quadratic cost function on quality production, the number of providers is a complicated function of the parameters of the model. The comparative statics of the explicit form for n^* in (83) have ambiguous signs, but the comparative statics can be solved from the general model by using the implicit function theorem. These features explain why analyzing the general model above is more useful than analyzing the explicit solutions for equilibria.

References

Abaluck, J. and Adams, A. 2017. What do consumers consider before they choose? Identification from asymmetric demand responses. NBER Working Papers 23566.

Abaluck, J. and Gruber, J. 2011. Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program. The American Economic Review 101, 1180–1210.

Amir, R., De Castro, L., and Koutsougeras, L. 2014. Free entry versus socially optimal entry. Journal of Economic Theory 154, 112–125.

Arrow, K. J. 1963. Uncertainty and the Welfare Economics of Medical Care. The American Economic Review 53 (5), 941–973.

Bardey, D., Canta, C., and Lozachmeur, J.-M. 2012. The regulation of health care providers' payments when horizontal and vertical differentiation matter. Journal of Health Economics 31 (5), 691–704.

Brekke, K., Cellini, R., Siciliani, L., and Straume, O. 2012. Competition in regulated markets with sluggish beliefs about quality. Journal of Economics & Management Strategy 21 (1), 131–178.

Brekke, K., Nuscheler, R., and Rune Straume, O. 2006. Quality and location choices under price regulation. Journal of Economics & Management Strategy 15 (1), 207–227.

Doyle, C., Lennox, L., and Bell, D. 2013. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ Open 3 (1).

Ellis, R. 1998. Creaming, skimping and dumping: provider competition on the intensive and extensive margins. Journal of Health Economics 17 (5), 537–555.

Ellis, R., Martins, B., and Miller, M. 2017. International Encyclopaedia of Public Health. In ed. by Heggenhougen, H. and Quah, S. Oxford: Academic Press. Chap. Provider payment methods and provider incentives, 133–143.

Gaynor, M., Ho, K., and Town, R. J. 2015. The industrial organization of healthcare markets. Journal of Economic Literature 53 (2), 235–284.

Handel, B. R. and Kolstad, J. T. 2015. Health insurance for "humans": Information frictions, plan choice, and consumer welfare. American Economic Review 105 (8), 2449–2500.

Kahneman, D., Wakker, P. P., and Sarin, R. 1997. Back to Bentham? Explorations of experienced utility. The Quarterly Journal of Economics 112 (2), 375–405.

Ma, C.-t. A. 1994. Health care payment systems: Cost and quality incentives. Journal of Economics & Management Strategy 3 (1), 93–112.

Ma, C.-t. A. and Mak, H. Y. 2018. Incentives in health care payment systems. Oxford Research Encyclopedia of Economics and Finance (Forthcoming).

Mak, H. Y. 2017. Managing imperfect competition by pay for performance and reference pricing. Journal of Health Economics.

Mankiw, N. G. and Whinston, M. D. 1986. Free entry and social inefficiency. The RAND Journal of Economics 17 (1), 48–58.

Newhouse, J. P. 1996. Reimbursing health plans and health providers: Efficiency in production versus selection. Journal of Economic Literature 34 (3), 1236–1263.

Competition in health care markets 2011. In Handbook of Health Economics. Ed. by Pauly, M. V., McGuire, T. G., and Barros, P. P. Vol. 2. Handbook of Health Economics. Elsevier, 499–637.

Salop, S. C. 1979. Monopolistic Competition with Outside Goods. The Bell Journal of Economics 10 (1), 141–156.

Spence, M. 1975. Monopoly, Quality, and Regulation. Bell Journal of Economics 6 (2), 417–429.

Tay, A. 2003. Assessing competition in hospital care markets: The importance of accounting for quality differentiation. The RAND Journal of Economics 34 (4), 786–814.

Waterson, M. 1993. Retail pharmacy in Melbourne: Actual and optimal densities. The Journal of Industrial Economics 41 (4), 403–419.

5 TEMPORARY AND PERSISTENT OVERWEIGHT AND LONG-TERM LABOR MARKET OUTCOMES

Abstract[*]
We examine the link between different durations of being overweight and long-term labor market outcomes. We use data on fraternal and identical twins born and raised in the same household. The data contain multiple weight measurements during early adulthood and information on subsequent earnings and employment over 20 years. When combined, these data enable an empirical strategy that controls for the family environment and genes shared by twins. We find that being persistently overweight during early adulthood is negatively associated with long-term earnings for both women and men. Persistently overweight women have approximately 30-40 percent and men have almost 45 percent smaller long-term earnings than those women and men who are never overweight. However, the underlying mechanism differs according to gender. For women, the association is driven by a decrease in labor market-attachment, whereas for men it is driven by lower annual earnings. Our results are robust to controlling for differences in initial health and physical conditions and educational achievement. Because our weight measurements are predetermined, our results are not biased by a contemporary correlation between weight and shocks to labor market outcomes. Our results highlight the importance of accounting for gender-specific heterogeneity and duration of being overweight when designing policy.

Keywords: Overweight, obesity, long-term labor market outcomes, labor market attachment.

[*] This chapter is joint work with Ari Hyytinen. We thank Anirban Basu, Randall Ellis, Jaakko Kaprio, Ching-to Albert Ma, Miikka Rokkanen, Benjamin Solow, Wenjia Zhu, and the participants of the ASHEcon 2014 in Los Angeles and the Boston University Department of Economics Health Economics reading group for their comments and suggestions. Laine received financial support from the Yrjö Jahnsson Foundation and the KAUTE Foundation. Data analysis has been carried out at Statistics Finland following its terms and conditions of confidentiality.

5.1 Introduction

Obesity is so common worldwide that it is considered an epidemic.[1] Obesity also has two major economic consequences. First, obesity raises medical costs (Cawley and Meyerhoefer 2012). Second, being overweight may affect an individual's labor market outcomes, such as earnings and employment (Averett 2011; Brunello and Beatrice 2007; Cawley 2015; Morris 2006, 2007; Paraponaris et al. 2005). Various policy interventions have been implemented aiming to mitigate the adverse effects of obesity on labor market outcomes, however with a little success.

This paper studies how different durations of being overweight are associated with long-term earnings and employment. Understanding these linkages is important for efficient policy design. For example, if the earnings penalty of being overweight depends on how long people have been overweight, policies that aim at reducing discrimination of overweight individuals are likely to have heterogeneous effects. The same applies to weight loss programs: if having been previously overweight is a determinant of the earnings penalty or of weaker labor market attachment, earlier interventions would be beneficial. As consequence, the importance for early obesity prevention programs, such as those promoted by the U.S. Department of Health and Human Services, becomes stronger if they can enhance long-term labor market outcomes in addition to the population's health by targeting children and adolescents. Despite the importance of the topic, very little is known about the relationship between timing and changes in overweight and long-term labor market outcomes, and even less is known about the mechanisms at work. This paper narrows this gap in the literature.

We use detailed administrative data on a large sample of fraternal and identical twins born and raised in the same household. Our empirical strategy exploits within-twin variation in these data to control for several key sources of omitted variable bias, the family environment, and genes shared by twins. Controlling for family environment is important because the shared home environment affects weight development (Gregory and Ruhm 2011; Segal and Allison 2002) and because there are family-level peer effects (Gwozdz et al. 2015). We also estimate models that fully control for genetic variation by focusing solely on identical twins. This analysis is motivated by the previous literature, which has documented that the genetically heritable component of weight is very high (Cawley 2004, 2015; Farooqi and O'Rahilly 2007). Compared to fraternal twins, who share slightly more than half of their genes, identical twins share all of their genes, and the rest of the differences in their genes are random (Zwijnenburg et al. 2010).[2] We can then fully remove the genetic effects by focusing on identical

1 In addition to the prevalence of adult obesity, also the prevalence in children and adolescents of obesity and being overweight has increased from 1975 to 2016, and if the post-2000 trends continue, child and adolescent obesity is expected to surpass being moderately and severely underweight by 2022 (Abarca-Gomez 2017).

2 We discuss factors causing variation in weights within twin pair in more detail Section

twins.[3]

Our data contain multiple weight measurements during the early adulthood of the twins and their subsequent earnings and employment over 20 years. Our main outcomes are long-term (total) earnings and long-term labor market attachment. They are constructed by using annual employment months and annual earnings, each of which is calculated as an average over a 20-year period that covers the prime working age of the twins. Based on the weight information we divide individuals into four mutually exclusive groups based on their overweight history, measuring whether individuals have been persistently overweight, whether they had lost or gained weight or whether they have never been overweight. Being able to distinguish individuals between these groups is important because the previous literature suggests that there is a great deal of heterogeneity in age-related temporal patterns in weight (see, for example, Livshits et al. 2012).

We find that being persistently overweight is adversely related with subsequent long-term earnings for both women and men. Persistently overweight women have from 30-40 percent smaller long-term earnings than women who are never overweight over the period during which the weight measurements took place. For men, we find that the corresponding earnings penalty is even larger, but only when genetics are fully controlled for. We also find that persistently overweight women work annually almost a month less than women who are never overweight. There is no such association for persistently overweight men. Instead, the persistently overweight men have lower monthly earnings, but again this is observed only when genetic differences are controlled for. We also find that subsequently overweight women have weaker labor market attachment as they work annually almost a month less than women who are never overweight.[4] There is no such relationship for subsequently overweight men. Our key findings are robust to controlling for remaining within-twin-pair variation in many important determinants of labor market outcomes, such as height, preexisting health conditions, and education.

Our results suggest that genetics shape the relation between overweight history and labor market outcomes more for men than for women. For men, the associations can be uncovered only when the genetics are fully controlled for. For women, the results are relatively robust across all specifications. This difference implies that the extent to which genetic factors are a source of omitted variable bias may differ across women and men. It also provides an explanation for the lack of consensus in the previous literature on the existence and magnitude of the earnings penalty for overweight men (Averett and Korenman 1996; Baum and Ford 2004; Behrman and Rosenzweig 2001; Cawley 2004; Sargent and Blanchflower 1994) and for why the evidence for men has been more mixed than that

5.3.1.

[3] This means that we can also control for differences in appearance, using the data on identical twins, see Jenq et al. (2015).

[4] Subsequently overweight individuals are those who were overweight or obese overweight during the last weight measurement period but were not overweight earlier.

for women.[5] Consistent with our findings, the recent genoeconomics literature has shown that the genetic heritability of weight is high and that the heritable component of long-term earnings is greater for men (Benjamin et al. 2012).

We also find that the underlying mechanisms behind the earnings penalty are gender-specific, too. For women, being the earnings penalty for being persistently overweight is related to weaker labor market attachment because the persistently overweight women work less than women who are not overweight. Because of this, and because the persistence in being overweight matters, the potential mechanism is likely to be related to something that erodes women's labor market attachment throughout their life, such as unobservable characteristics (Falkner et al. 2001; Kristjansson et al. 2010; Sabia and Rees 2012) and skill formation earlier in life (Florin et al. 2011; Heckman 2006; Mobius and Rosenblat 2006). For overweight men, the mechanism may be related to something that erodes men's earnings capacity in the labor market but not their labor market attachment.

We complement the limited existing literature studying whether the timing of overweight matters for the earnings penalty (Chen 2012; Han et al. 2011; Pinkston 2017) in several fundamental ways. First, our unique data include longitudinal information on both outcomes, weight measures, and information on twins. The previous literature has only used cross-sectional measures for outcome(s), weight measures, or both: Han et al. (2011), Chen (2012), and Pinkston (2017) use the hourly rate of pay of the most current job as a key outcome measure. We focus instead on much longer-term labor market outcomes that are calculated over a period of 20 years. The focus on long-term labor market outcomes is an important extension because it allows us to capture the cumulative labor market effects of being overweight. Our measurement for the timing of being overweight also differs from Han et al. (2011), Chen (2012), and Pinkston (2017). In Han et al. (2011), weight is measured using data on BMI in late teenage years and current BMI, whereas in Chen (2012) and Pinkston (2017) weight is measured in early or late adulthood. We instead use three different weight measurements over a period that covers 15 years. It is the combination of the three long-term outcome measures and access to the twin data that makes our analysis unique.[6] This combination of longitudinal information on outcomes, independent variables (weight measures) and zygosity of twins provides us with a novel way of studying the mechanisms behind the weight penalty.

In addition to being able to study the link of being overweight earlier or later in life on long-term labor market outcomes, our approach allows us to study the association of being persistently overweight on long-term labor market outcomes. To the best of our knowledge this is the first paper to study the link between persistent overweight and labor market outcomes. Unlike previous studies (Han et al. 2011; Chen 2012; Pinkston 2017), we are able to show that what matters for long-term labor market outcomes is being chronically overweight,

5 For a summary of the recent literature, see Cawley (2015).

6 The first weight measurement took place when the individuals were, on average, 23 years old and the last when they were, on average, 38 years old.

that the mechanisms at work differ by gender, and that the association of being chronically overweight with adverse long-term outcomes is not driven by genetic differences. If being overweight reduces people's lifetime earnings as compared to it only reducing current earnings, the implications for lifetime consumption possibilities and for health policy are quite different.

A challenge in the literature studying the economic consequences of being overweight and obese is to find plausibly exogenous variation in weight that allows one to identify the effect of being overweight or obese. Prior literature has used various methods and instruments to solve this problem. One of the more current approaches is to use genetic risk scores as an instrument (Böckerman et al. 2018). While we find the approach of using genetic risk scores novel and valuable, the approach cannot be used here to address the specific research question that we are interested in. First, as far as we are aware, medical science and behavioral genetics have not identified genes that would predict individuals' weight in different parts of their life-cycle or, relatedly, the duration of being overweight. Second, even if such instruments became available, it is not clear that they would satisfy the required exclusion restriction for instrumental variable estimation in our context.

The remainder of this paper is structured as follows. Section 5.2 describes our data and descriptive statistics. We describe our econometric approach in Section 5.3. We report and analyze our main results in Section 5.4. Section 5.5 discusses our robustness checks, such as alternative definitions of BMI-thresholds, outcome measures, characterization of weight history, and additional controls. Section 5.6 provides a discussion and policy suggestions, and Section 5.7 concludes.

5.2 Data and variables

5.2.1 Data

We use two unique and detailed data sets. First is the Finnish Longitudinal Employer-Employee Data (FLEED), which is an annual panel that covers the working-age population in Finland. The data available to us cover the period from 1990 to 2009. FLEED is constructed from several different administrative registers, including annual information on earnings and other income such as social benefits, employment, and education. An advantage of FLEED is that our data on earnings and employment data do not suffer from the usual survey data problems, such as recall error or over- or underreporting. Also, the earnings data are not top-coded. We use FLEED to create our measures for long-term labor market outcomes.[7]

The second data set is the Older Finnish Twin Cohort Study, which was

[7] The data set is maintained by Statistics Finland. For a more detailed description of FLEED, see Hyytinen et al. (2013).

launched in 1974.[8] The Older Finnish Twin Cohort Study is a survey that covers all Finnish twin pairs born before 1958 and for which both co-twins were alive in 1975.[9] The survey was carried out in 1975, 1981, and 1990.[10] We use weight and height information from the three twin surveys to derive measures for Body Mass Index (BMI). We use BMI to construct our main regressors of interest.

The Twin Survey data were linked to FLEED using personal identifiers.[11] We restrict our sample to same-sex twin pairs born after 1944 but before 1958. We use this restriction so that our main outcome variables are measured over a period that covers the primary working age of the individuals in our data. The individuals were from 32 to 46 years old in 1990 and from 51 to 65 years old in 2009. The estimation sample consists of 2,364 women and 1,564 men.[12] This corresponds to 1,182 female twin pairs of which 747 are fraternal twin pairs and 435 identical twin pairs. And it corresponds also to 782 male twin pairs of which 515 are fraternal twin pairs and 267 identical twin pairs.

5.2.2 Variables

We construct three long-term labor market outcome variables using FLEED. First is a proxy for lifetime earnings, *Long-term earnings*. We calculate it as the natural logarithm of the average annual earnings over 1990–2009.[13] Our second outcome variable is the individual's labor market attachment, *Average employment*. We calculate it as the average annual employment months over 1990 –2009. Our third outcome variable is the average monthly earnings over 1990–2009, *Monthly earnings*. It is the natural logarithm of the ratio of average annual earnings and average annual employment months over 1990–2009.

The focus on long-term labor market outcomes is an important extension to the previous literature because it allows us to capture the cumulative labor market effects of being overweight. If investments in skills and capabilities are dynamic complements over time (Heckman and Mosso 2014), and if being over-

8 Record linkages of the cohort study data conform to the Data Protection Act and were originally approved by the ethical committee of the Department of Public Health, University of Helsinki. Statistics Finland has accepted the record linkages used for the matched data used in this paper. All the data work of this paper was carried out at Statistics Finland, following its terms and conditions of confidentiality. For a detailed description of the Older Finnish Twin Cohort Study of the Department of Public Health at the University of Helsinki, see Kaprio et al. (1979).

9 Twin zygosity was determined by a questionnaire method and was subsequently validated by blood markers (Kaprio et al. 1979)

10 The twin pairs were originally selected from the Central Population Registry of Finland, and in the first twin survey of 1975, questionnaires were mailed to all pairs. The first survey had two follow-up studies that were carried out in 1981 and 1990. The response rate in the survey of 1975 was 89 percent, in 1981 the response rate was 84 percent, and in 1990 the response rate was 77 percent. Unlike the first two surveys, the 1990 survey was sent only to those twins who were born after 1930.

11 For a description of the combined data set see, for example, Hyytinen et al. (2013).

12 Over representation of women common in survey based twin data (Silventoinen et al. 2015, Figure 2).

13 By earnings, we mean the sum of wages, salaries, and entrepreneurial income.

weight affects schooling and early career choices (Han et al. 2011), skill formation, and learning-by-doing, then these cumulative effects can be observed only in data sets like ours. In addition to the pure long-term earnings and monthly earnings, we also study the relation of being overweight on long-term average employment. This measure reflects genuine labor market attachment better than cross-sectional measures of an individual's labor market status, such as temporary unemployment spells. This allows us to distinguish if the relation of being overweight is caused by a direct effect on earnings or weaker labor market attachment. We test the robustness of our results by re-estimating our models using an alternative measure for long-term earnings and monthly earnings. The results of these robustness checks are collected in Section 5.4.

To construct measures for an individual's overweight history, we determine how their weight has developed over time. For this, we use BMI (kg/m^2) as a proxy of excess body fat. Our BMI measures are based on self-reported weight (in kilograms) and height (in meters) from the three twin surveys of 1975, 1981, and 1990. We classify an individual as being overweight if her/his $BMI \geq 25.00$.[14] This threshold is the same for both sexes. Our threshold choice is in line with Gregory and Ruhm (2011) and other prior work that suggests that the BMI range relevant for adverse labor market outcomes may be below the obesity threshold $BMI \geq 30.00$.

We divide individuals into four groups based on their overweight history by using the three different weight measurements from 1975, 1981, 1990. The four groups are called *Persistently overweight*, *Previously overweight*, *Subsequently overweight*, and *Never overweight*. We call an individual persistently overweight, $BMI_{ij}^{persistently} = 1$, if the individual was overweight in either 1975 or 1981 (or both) and was also overweight in 1990. We call an individual previously overweight, $BMI_{ij}^{previously} = 1$, if the individual was overweight in 1975 or 1981 but was not overweight in 1990. We call an individual subsequently overweight, $BMI_{ij}^{subsequently} = 1$, if the individual was overweight during the last survey year 1990 but was not overweight in 1975 or 1981. Last, we call an individual never overweight, $BMI_{ij}^{never} = 1$, if the individual was not overweight in any of the survey years 1975, 1981, and 1990. We study the robustness of our results from models that use alternative BMI classifications and BMI thresholds (see Section 5.2. for more details).

BMI is a measure that adjusts body weight for height. A valid critique against BMI as a measure of body fat and being overweight is that it does not take age, sex, or body composition into account (see, for example, Burkhauser and Cawley 2008; Johansson et al. 2009; Rothman 2008). For instance, because muscles weigh more than fat, BMI can underestimate the amount of body fat of individuals with low muscle mass. Similarly, it can overestimate the amount of

---

14 See, for instance, the WHO Global Database on Body Mass Index. http://apps.who.int/bmi/index.jsp?introPage=intro_3.html. We have also re-estimated our models by using a different threshold specification that divides BMI into 3-year and gender-specific terciles. We discuss these results below in Section 5.2.

body fat on lean individuals with high muscle mass.

This standard critique against BMI is less of a concern to us because our sample consists of twins born and raised in the same household. This is beneficial for two reasons. First, the body composition between twins can be assumed to be more similar than for two randomly chosen individuals. Even though body composition can differ between siblings (Price and Swigert 2012), two randomly chosen individuals differ even more due to between-family variation in environment and genes. Second, the body composition of identical twins is as similar as it can biologically be because in addition to their home environment, identical twins share almost all genes (Zwijnenburg et al. 2010).

5.2.3 Summary statistics

Table 1 reports summary statistics for the outcome variables, the overweight history categories, and the control variables for the full sample of twins. There are several differences between men and women. The mean of long-term earnings is 9.37 for women, corresponding to approximately 14,900 euros per year. The corresponding number for men is 9.70, corresponding to approximately 21,308 euros. Over our sample period 1990-2009 women worked 7.9 months per year, and their monthly mean earnings were 2,300 euros. The corresponding numbers for men are larger: men worked 8.2 months per year, and their monthly mean earnings were 2,600 euros.

TABLE 1 Descriptive statistics (means).

	Women	Men
Panel A: Dependent variables (1990-2009)		
Long-term earnings (log)	9.37	9.70
Average employment months (months)	7.90	8.19
Monthly earnings (log)	7.42	7.71
Panel B: BMI-category dummies		
Never overweight	0.75	0.56
Persistently overweight	0.10	0.24
Previously overweight	0.02	0.03
Subsequently overweight	0.13	0.17
Panel C: Control variables		
Age in 1990 (years)	38.7	39.2
Height in 1975 (cm)	163.5	176.6
Schooling (years)	12.2	12.2
Number of diseases 1975	0.50	0.39
Number of observations	2364	1564

Note. Variables in Panel A in addition to age and schooling variables in Panel C are from FLEED. Variables in Panel B in addition to the height, number of diseases, and employment information in 1975 are from the Older Finnish Twin Cohort Study.

Both sexes were on average 39 years old in 1990. Our education variable refers to the number of years in school, which is based on information on achieved degrees and standard degree times. Both sexes had, on average, 12 years of schooling in 1990. The table also shows that men were slightly healthier than women as measured by the number of diseases they reported in the 1975 survey. Additionally, Table A2 in Appendix 5.A summarizes the number of observations and the corresponding means of the absolute values of the twin-differenced variables for the fraternal and identical twins. The means of the absolute values of the twin-differenced variables indicate that there is quite a bit of variation in the twin-differenced data as well. For example, the average (absolute) BMI difference of persistently overweight identical twins is as large as 0.19 for men and 0.09 women.[15] Table 1 also shows that 10 percent of women and nearly 24 percent of men were persistently overweight during the sample period. Also, 13 percent of women and 17 percent of men who were not overweight initially were overweight in 1990 when their weight was measured the last time. The percentage of those who had previously been overweight but were no longer overweight in 1990 are small for both women and men, approximately 2-3 percent.

Table 2 displays the means of BMI measured at different points in time, age in 1990, and the outcome variables by the overweight history categories, as well as p-values from F-tests for equality of the group means. The table depicts a few noteworthy patterns. First, average weight increases over time. For example, the average BMI of persistently overweight women increases from 28.0 in 1975 to 32.0 in 1990. For men, the corresponding increase is from 27.0 in 1975 to 28.9 in 1990. Second, persistently overweight individuals are, on average, somewhat older than the rest of the individuals. Third, average long-term earnings are lowest for women and men who are either persistently or subsequently overweight. Labor market attachment, as measured by average annual months of employment, is also weakest for those who were persistently overweight. As the p-values of the F-tests show, the differences in the means of the various overweight categories are all statistically significant.

15 Figure A1 in our Appendix 5.A depicts the Kernel density estimates of the natural logarithm of long-term earnings, average employment months, and monthly wages conditional on different obesity history categories for women (left panel) and men (right panel) separately. The figure shows that the individuals who are never overweight work more.

TABLE 2 Means in different overweight history categories.

	Never overweight	Persistently overweight	Subsequently overweight	Previously overweight	p-value of F-test
			Panel A: Women		
BMI					
BMI in 1975	20.1	25.4	21.8	24.4	< 0.01
BMI in 1981	20.6	28.0	22.9	25.0	< 0.01
BMI in 1990	21.6	32.0	26.9	23.3	< 0.01
Age in 1990	38.5	39.8	39.1	39.2	< 0.01
Dependent variables (1990-2009)					
Long-term earnings (log)	9.42	9.12	9.30	9.06	< 0.01
Average employment months (months)	8.02	7.40	7.63	7.73	< 0.01
Monthly earnings (log)	7.46	7.26	7.39	7.22	< 0.01
Number of observations	1766	241	306	51	
			Panel B: Men		
BMI					
BMI in 1975	21.2	25.3	22.4	24.3	< 0.01
BMI in 1981	21.8	26.9	23.5	25.4	< 0.01
BMI in 1990	22.6	28.2	26.4	24.2	< 0.01
Age in 1990	38.7	40.6	38.7	39.9	< 0.01
Dependent variables (1990-2009)					
Long-term earnings (log)	9.72	9.63	9.77	9.60	< 0.01
Average employment months (months)	8.31	7.89	8.29	7.90	< 0.01
Monthly earnings (log)	7.71	7.68	7.73	7.65	< 0.01
Number of observations	874	369	272	49	

Note. The first four columns report the average characteristics for the four overweight history categories. The last column reports p-values from F-tests for equality of the group means.

5.3 Empirical framework

5.3.1 Empirical strategy

Our main analysis uses the following econometric model:

$$Y_{ij} = \beta_1 BMI_{ij}^{persistently} + \beta_2 BMI_{ij}^{previously} + \beta_3 BMI_{ij}^{subsequently} + \gamma' Z_{ij} + \delta_j + \epsilon_{ij}, \qquad (1),$$

where Y_{ij} is one of our three outcome variables, $Long\text{-}term\ earnings_{ij}$, $Average\ employment_{ij}$ and $Monthly\ earnings_{ij}$, of twin i in pair j, and where $BMI_{ij}^{persistently}$, $BMI_{ij}^{previously}$, and $BMI_{ij}^{subsequently}$ are our variables for persistent, previous, and subsequently overweight. The omitted overweight history category refers to those who are never overweight, and Z_{ij} includes the observed control variables. δ_j denotes the unobserved family environment and genes shared by twins, and ϵ_{ij} is the individual specific error term.

We use two estimation strategies. Our first strategy ignores δ_j and treats it as part of the error term. This amounts to estimating the model in levels using variation both within and across twins (OLS regression). The consistency of this estimator requires that BMI be uncorrelated with both δ_j and ϵ_{ij}. We collect the results of this estimation strategy in Section 5.4.1.

Our second estimation strategy controls for δ_j by including twin fixed effects in the model. This amounts to estimating the model in within-twin differences and is our main identification strategy. This strategy controls for family environment and genes shared by twins. The consistency of this model does not require that BMI be uncorrelated with δ_j. We begin by estimating the twin-differenced model using the full sample, which includes both fraternal and identical twins and collect the results of this estimation strategy in Section 5.4.2.

We also estimate the twin-differenced model using a sub-sample of identical twins. This specification allows us to study how unmeasured genetic attributes affect the estimates. This is important because the heritable component of weight is very high (Cawley 2004, 2015; Farooqi and O'Rahilly 2007) and because labor market outcomes are also known to have such a component (Benjamin et al. 2012). The results of using the second estimation strategy using the sub-sample of identical twins are collected in Section 5.4.3.

We consider three different regression specifications. The first specification is equation (1) without control variables, Z_{ij}. Second, we use control variables from both the FLEED and twin survey data. In our main specification, we include as control variables exogenous or predetermined determinants of our long-term labor market outcomes. These control variables are age, age squared, and height.[16] Our key findings are robust to controlling for remaining within-twin-pair variation in many important determinants of labor market outcomes, such as height, pre-existing health conditions, and education.

In our third specification, we also include as control variables education (years of schooling in 1990) and health (number of diseases in 1975). One motivation to add health is the well-documented connection between being overweight and poor health outcomes. Moreover, Han et al. (2011) provide evidence that being obese in late teenage years reduces the amount of acquired schooling and affects occupation outcomes. However, it is not at all clear that health and education should be controlled because we are interested in estimating the total lifetime earnings effect of being overweight. These additional controls may themselves be affected by being overweight, thus giving rise to the bad control problem (Angrist and Pischke 2009). Nevertheless, as a robustness check we include these commonly used controls to see whether they can explain away some of the effect of being overweight.

In sum, we have three regression specifications for each of the three outcome variables. This totals 9 models. Our baseline results therefore consist of 27 estimations because each of these nine models is estimated in levels (approach 1). We then use the two versions of the fixed effects regression (approach 2). We report standard errors clustered by twin pairs for models estimated in levels and heteroskedasticity-robust standard errors in twin-differences.

Previous papers have used various instruments such as sibling-instruments (Cawley 2004; Cawley and Meyerhoefer 2012; Kline and Tobias 2008; Lindeboom et al. 2010), lagged-weights (Averett and Korenman 1996; Baum and Ford 2004;

16 Recall that our sample includes fraternal and identical twins and within twin-pair variation also in height.

Cawley 2004; Sargent and Blanchflower 1994), the prevalence of obesity (Morris 2007), and genetic risk scores (Böckerman et al. 2018) when identifying the association of being obeseity and labor market outcomes. Unfortunately this approach cannot be used to address the specific research question that we are interested in. In particular, we are unaware of variables that would predict individuals' weight in different parts of their life-cycle or, relatedly, the duration of being overweight. Moreover, even if such variables became available, it is not clear that they would satisfy the required exclusion restriction for instrumental variable estimation in our context.

Our approach is complementary and use panel data on twins, combined with detailed longitudinal administrative data to estimate the association of being overweight and labor market outcomes. There are a few advantages of our approach. First is that we can use within-twin variation to control for several key sources of omitted variable bias, namely, the family environment and genes shared by twins, as well as for factors that may make two identical twins different.[17] Another benefit of using within-twin-pair variation among the identical twins is that it removes concern over the effect of being overweight being driven by an omitted variable related to individuals' appearance. We can largely rule out this possibility because identical twins are biologically very similar (this has been emphasized by Gregory and Ruhm 2011).

A few possible concerns with our identification strategy should be pointed out. First is that one twin's weight may deviate from the co-twin's weight for reasons that are correlated with the regression error term. For example, it is possible that because of a physical injury or a health shock one of the twins gains weight and participates and earns less in the labor market. While we cannot completely rule out such shocks, we would like to argue that they are unlikely to bias our estimates for the following reasons. First, our measures of BMI are predetermined, as they are measured much earlier than the labor market outcomes. This reduces the concern that the (contemporary) error term in the labor market outcome regressions is correlated with the past weight measures.

In particular, a physical or health shock can bias for example the coefficient of being persistently overweight only if it happened before the weight measurements took place and if it still continues to have a labor market impact much later over the 20-year period from 1990 to 2009. Our results are also robust to controlling for within-pair differences in initial health and physical conditions, both measured in 1975. More importantly, the results are also robust to within-pair differences in education, which is a key determinant of permanent earnings. This means that whatever the shock that affects both weight and labor market outcomes of one twin, but not the other, it is a source of bias only if its long-term

---

17 Our strategy follows the broad literature that has used within-sibling variation to study, for instance, the effects of education on economic outcomes in different contexts (Aaronson 1998; Abramitzky et al. 2012; Altonji and Dunn 1996; Griliches 1979; Sacerdote 2007). For literature specifically using twin-pairs to estimate the returns of education, see Behrman and Rosenzweig (2002) and Behrman et al. (1996).

labor market effects cannot be proxied by differences in education.[18] Another concern relates to the some of the BMI measures were taken place while some of the individuals were already at their prime working age. This means that we cannot rule out the possibility of reverse causality.

Last, previous literature shows that there is a strong influence of genetic factors on BMI: almost 80 percent of BMI is heritable. Rest of the variation variation in weights within twin pair could be explained by behavioral factors such as nutrition, physical activity, and smoking thus leading to a concern of our approach not being able to remove the all potential sources of omitted variable bias. While we acknowledge these concern, we point out that this is a less of a concern to us because instead of causal claims our focus specifically is to focus the pinning down the association between different durations of being overweight and long-term labor market outcomes.

5.3.2 Potential mechanisms

Several mechanisms, such as lower productivity and various forms of labor market discrimination, have been proposed as explanations for the negative relation of being overweight on labor market outcomes (see, for example, Cawley 2004). Our econometric framework allows us to explore the underlying economic mechanisms behind the adverse wage effects of being overweight in two ways.

The first is related to our choice of outcome measures, which are long-term earnings, average employment months, and monthly earnings. The first of these measures allows us to evaluate the overall association of being overweight on the individual's lifetime earnings. If there is such an association, the latter two outcome measures can be used to explore this relation in more detail. For instance, lower long-term earnings might be caused by a labor supply effect, which refers to weaker labor market attachment. It might be also caused by a wage penalty effect, which refers to lower earnings per amount worked. Lower long-term earnings might also be a combination of these two.

The second approach relies on our ability to measure individuals' weight at different points in their early adulthood. This allows us to characterize an individual's weight history, that is whether the individual has been persistently, previously, or subsequently overweight. This is important because many of the mechanisms proposed by the previous literature are likely to be more crucial at certain points in life. This makes testing of the relative importance of the types of individuals' overweight histories relevant.

One suggested mechanism is related to general skill formation in childhood (especially skills that are learned outside of school). Being overweight as a child or in early adulthood can cause underdeveloped social skills, lower self-esteem, and poor communication skills. It can therefore negatively affect an individual's

18 Having children has been suggested to reduce labor supply, especially for women (see, for example, Cools et al. 2017). Unfortunately, we do not have information on children in our data. However, the number of children might be a bad control if being overweight affects also the number of children.

career path and cause lower wages in the future (Florin et al. 2011; Heckman 2006; Mobius and Rosenblat 2006). This kind of mechanism is likely to work throughout an individual's life. Another mechanism that is likely to work throughout an individual's life is whether being overweight is a proxy for some unobservable characteristics. For instance, overweight individuals may have high discount rates and therefore invest less in education, health, and weight control (Falkner et al. 2001; Kristjansson et al. 2010; Sabia and Rees 2012). These kinds of cumulative effects of being overweight on labor market outcomes are captured by persistent and previous overweight. If the underlying mechanism is related to skill formation or permanent unobservable characteristics, we should see a non-negligible negative effect on long-term labor market outcomes for those who have been persistently or previously overweight.

Other mechanisms are more relevant in adulthood than in childhood. One is labor market discrimination, which means that promotions and salary raises are not granted as often to workers who are currently less attractive physically (Mobius and Rosenblat 2006; Puhl and Brownell 2001), or employers discriminate against overweight individuals based on their appearance when choosing which job applicants to interview and hire (Gregory and Ruhm 2011; Han et al. 2009; Rooth 2009). Previous literature suggests discrimination against obese workers is worse against women (Caliendo and Lee 2013).

Another potential mechanism that is more relevant in adulthood is that lower wages mirror lower productivity. Subsequently overweight workers may, for example, be absent from work more often or less handy in performing certain types of physically demanding tasks. Their lower productivity may show up as lower wages. If the underlying mechanism is related more to labor market discrimination or lower productivity, we should see a non-negligible negative effect of being subsequently overweight on long-term labor market outcomes.[19]

The estimated coefficients for being persistently overweight (β_1), having been previously overweight (β_2), and being subsequently overweight (β_3) enable testing the relative importance of various weight histories. At least the following five comparisons are interesting. First, we can test whether being overweight matters at all by testing the null hypothesis $\beta_1 = \beta_2 = \beta_3 = 0$. Second, if the null hypothesis of $\beta_1 = \beta_2 = \beta_3$ is rejected, the effects of being overweight in different parts of life are different. Third, we can test whether the effect of being persistently overweight is the same as the effect of having been previously overweight. If the null hypothesis $\beta_1 = \beta_2$ is rejected, we can infer that being subsequently overweight matters more than previously being overweight for those who have been overweight earlier in life. Third, we can test whether the effect of persistently being overweight is the same as that of being subsequently overweight. If

[19] It has also been suggested that firms impose a wage penalty for overweight individuals to compensate for their higher medical costs on their employer-provided health insurance (Bhattacharya and Bundorf 2009). This mechanism is more relevant in countries in which firms are required to provide employer-sponsored health insurance. It is unlikely that this mechanism is empirically relevant in Finland because health insurance in Finland is largely publicly funded. Finnish employers have no or weaker incentives to internalize the medical costs of obesity.

the null hypothesis $\beta_1 = \beta_3$ is rejected, it is not. Fourth, we can test whether the effect of being previously overweight is the same as that of being subsequently overweight. If the null hypothesis $\beta_2 = \beta_3$ is rejected, it is not.

Finally, we can test whether being overweight in the past interacts with being subsequently overweight. If the null hypothesis $\beta_1 = \beta_2 + \beta_3$ is rejected, the effect of being persistently overweight is larger (or smaller) than the sum of the effects of having been overweight in the past and of being subsequently overweight. To the best of our knowledge, this interaction has not been studied before. For instance, Chen (2012) was unable to study this because her data included information only about the previous and current weight. The effect of being persistently overweight could be larger than just the sum of being previously and subsequently overweight if the adverse effects accumulate or increase over time. It could also be smaller if those who are persistently overweight learn to take compensatory measures to alleviate the adverse labor market consequences of being overweight.

5.4 Results

5.4.1 Baseline results: OLS regressions

Our first estimation approach uses model (1) in levels and variation both within and across twins (specification 1). Table 3 displays the results for three outcome variables and for the three different regression specifications, which totals nine estimated models for each gender. Panel A of the table reports the results for women and panel A the results for men. The lower parts of both panels summarize the p-values for the test of the relative importance of various patterns of overweight history on the labor market outcomes.

We find different relations of being overweight for women and men. Our results show that persistently overweight women have approximately 20 to 30 percent smaller long-term earnings than women who are never overweight.[20] Persistently overweight women also work 0.4 to 0.6 months less and have 13 to 19 percent smaller monthly earnings than women who have never been overweight. The OLS coefficients for women become somewhat smaller (in absolute values) when education and health are included as controls. For men, we find no similar effects.

The results for transitory obesity measures, previously overweight and subsequently overweight, are mixed. The coefficients are consistently negative for women, but the estimated relations are no longer statistically significant at the 5 percent level when education and health are included as controls. For men, the coefficients are smaller and are never significant in regressors that include con-

20 Because our dependent variable *Long-term earnings* is a logarithmic transformation of the long-term earnings, the coefficients of the BMI variables can be interpreted in terms of percentage changes.

trols.

The tests reported in the lower parts of the panels show two interesting findings for women. First, being overweight matters for women, as the joint null hypothesis $\beta_1 = \beta_2 = \beta_3 = 0$ is rejected for all outcome variables. We also reject the hypothesis that the relation between long-term earnings and monthly earnings of being persistently overweight is the same as the corresponding relation of being subsequently overweight. This finding provides support for the view that the lifetime earnings penalty is greater for persistently overweight women than for subsequently overweight women. We do not want to overemphasize these OLS results. Previous studies have shown that family environment and genetic effects are potential sources of omitted variable bias because of their association with variation in overweight and long-term labor market outcomes. Another critique concerns BMI, which is the most commonly used measure of obesity and being overweight. Using BMI as an overweight measure can bias the estimates because it does not take into account body composition (that is the relative share of fat, bone, water, and muscle in human bodies; see Burkhauser and Cawley 2008; Johansson et al. 2009). Fortunately, we can address some of these concerns because our sample consists of twins born and raised in the same household, and their body composition and appearance are arguably more similar than that of two randomly chosen individuals. We therefore next exploit variation within twins to eliminate all potential bias coming from unobserved family environment and body composition.

5.4.2 The role of family environment

Our second approach is to estimate models with twin fixed effects (specification 2), using the full samples of women and men, including both fraternal and identical twins. The results of these estimations are reported in Table 4, separately for women (Panel A) and men (Panel B). The p-values corresponding to the tests of the relative importance of weight history patterns are summarized in the lower parts of the two panels. The table shows several interesting findings. First, we find more and stronger associations for women after the effect of family environment and genes shared by twins are controlled for. Persistently overweight women have approximately 30 percent smaller long-term earnings than women who are never overweight. Persistently overweight women also work 0.8 months less and have $15 - 17$ percent smaller monthly earnings than women who are never overweight. For men, there is no such relation in regard to being being persistently overweight.

Also, for subsequently overweight women, the relation of being overweight on average employment months becomes stronger after the family environment and genes shared by twin siblings are controlled for.[21] There is no such relation for subsequently overweight men. How do previously overweight individuals who have been able to lose weight differ from the others? We do not emphasize

[21] Fraternal twins share slightly more than half of their genes, and identical twins share almost all genes.

TABLE 3 OLS regressions, full sample for women and men separately.

	Long-term earnings			Average employment months			Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Women				
BMI-categories									
Persistently overweight (β_1)	-0.306***	-0.277***	-0.207***	-0.628***	-0.521***	-0.443**	-0.192***	-0.188***	-0.130***
	(0.067)	(0.065)	(0.062)	(0.195)	(0.192)	(0.188)	(0.039)	(0.039)	(0.036)
Previously overweight (β_2)	-0.365**	-0.359**	-0.298	-0.289	-0.297	-0.236	-0.233**	-0.228**	-0.176
	(0.171)	(0.169)	(0.178)	(0.420)	(0.414)	(0.421)	(0.109)	(0.109)	(0.113)
Subsequently overweight (β_3)	-0.125**	-0.106**	-0.057	-0.389**	-0.332	-0.289	-0.063**	-0.058	-0.014
	(0.051)	(0.051)	(0.050)	(0.168)	(0.169)	(0.169)	(0.031)	(0.031)	(0.029)
Controls									
Age	N	Y	Y	N	Y	Y	N	Y	Y
Height	N	Y	Y	N	Y	Y	N	Y	Y
Education	N	N	Y	N	N	Y	N	N	Y
Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.000	0.000	0.003	0.003	0.017	0.055	0.000	0.000	0.002
Joint test of controls	-	0.001	0.000	-	0.000	0.000	-	0.000	0.000
$\beta_1 = \beta_2 = \beta_3$	0.043	0.051	0.083	0.541	0.701	0.774	0.009	0.008	0.016
$\beta_1 = \beta_2 + \beta_3$	0.329	0.313	0.445	0.915	0.818	0.864	0.382	0.411	0.619
$\beta_1 = \beta_2$	0.745	0.646	0.624	0.444	0.610	0.640	0.721	0.727	0.697
$\beta_1 = \beta_3$	0.022	0.030	0.047	0.325	0.432	0.517	0.004	0.004	0.008
$\beta_2 = \beta_3$	0.170	0.146	0.183	0.820	0.937	0.904	0.126	0.127	0.162
Number of observations	2364	2364	2364	2364	2364	2364	2364	2364	2364
					Panel B: Men				
BMI-categories									
Persistently overweight (β_1)	-0.097	-0.077	0.036	-0.418**	-0.198	-0.058	-0.027	-0.044	0.038
	(0.062)	(0.061)	(0.058)	(0.170)	(0.162)	(0.159)	(0.037)	(0.036)	(0.034)
Previously overweight (β_2)	-0.122	-0.108	-0.015	-0.402	-0.288	-0.189	-0.064	-0.067	0.005
	(0.134)	(0.128)	(0.122)	(0.421)	(0.380)	(0.373)	(0.076)	(0.077)	(0.072)
Subsequently overweight (β_3)	0.048	0.048	0.060	-0.015	-0.011	0.008	0.018	0.017	0.024
	(0.060)	(0.060)	(0.055)	(0.169)	(0.168)	(0.165)	(0.038)	(0.038)	(0.033)
Controls									
Age	N	Y	Y	N	Y	Y	N	Y	Y
Height	N	Y	Y	N	Y	Y	N	Y	Y
Education	N	N	Y	N	N	Y	N	N	Y
Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.161	0.284	0.718	0.072	0.577	0.943	0.644	0.443	0.697
Joint test of controls	-	0.000	0.000	-	0.000	0.000	-	0.000	0.000
$\beta_1 = \beta_2 = \beta_3$	0.088	0.161	0.813	0.122	0.571	0.859	0.460	0.319	0.879
$\beta_1 = \beta_2 + \beta_3$	0.878	0.906	0.947	0.998	0.812	0.768	0.828	0.941	0.913
$\beta_1 = \beta_2$	0.856	0.812	0.684	0.970	0.817	0.733	0.642	0.763	0.656
$\beta_1 = \beta_3$	0.037	0.075	0.707	0.045	0.340	0.733	0.305	0.169	0.729
$\beta_2 = \beta_3$	0.219	0.241	0.548	0.374	0.485	0.612	0.314	0.297	0.797
Number of observations	1564	1564	1564	1564	1564	1564	1564	1564	1564

Notes: Full sample includes both fraternal and identical twins. Dependent variables are lifetime earnings (log), average employment months (months), and monthly earnings (log), calculated as the average over 1990 − 2009. Explanatory variables are persistently overweight, previously overweight, and subsequently overweight (omitted category = *never overweight*), determined by BMI measured in 1975, 1981, and 1990 with BMI threshold > 25 for being overweight. The control variables are age and age squared in 1990, height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. Standard errors are in parentheses (clustered by twin pair). **$p < 0.05$, ***$p < 0.01$.

the results for the previously overweight. Our estimated relations of having been previously overweight on long-term labor market outcomes are mostly insignificant and not very robust. This is because the number of female and male twin pairs in which one, but not both, of the twins would belong to the overweight history category of previously overweight individuals is very small (see also Table 1). Therefore, the effect of losing weight is more difficult to quantify. One reason for this is that the number of individuals who were overweight in the past and who become a normal weight at some point later in life is relatively small. In our data, only 3 percent of individuals could do so.

Third, we find that for women, the p-values reported in the lower part of Panel A suggest that being overweight matters for long-term earnings and labor market attachment. For these two outcome variables, the joint null hypothesis $\beta_1 = \beta_2 = \beta_3 = 0$ is rejected at better than the 5 percent level.

5.4.3 The role of genetics

We then estimate the model with twin fixed effects (specification 2), using the sample to identical twins only. This allows us to explore the importance of unobserved genetic effects: twin differencing removes the genetic effects because identical twins share almost all their genes.

Estimating the model with twin fixed effects using the full sample including both fraternal and identical twins removes the for family environment and genes shared by twins. While all twins in our sample have the same family environment, there is variation in the extent to which they share the same genes. Fraternal twins share slightly more than half of their genes. Previous studies suggest that genetics may be associated with variation in obesity (Norton and Han 2008) and long-term income (Benjamin et al. 2012; Jäntti and Lindahl 2012). This means that using the full sample that includes both fraternal and identical twins may not be enough to eliminate omitted variable bias.

Table 5 summarizes the within-twin regression results for our sample of identical twins. The main finding arising from this approach is that genes seem to affect labor market outcomes of men more than those of women. The results for women are in line with our previous full-twin sample results. We find that persistently overweight identical twin women have approximately 40 percent smaller long-term earnings and work approximately 1.1 months less than women who are never overweight. This suggests that the wage penalty of overweight on long-term earnings is mainly driven by persistently overweight women working less. The only difference compared to our findings above is that the association between persistently being overweight and on women's monthly earnings is no longer statistically significant.

We obtain new results for men after family environment and genetics are controlled for. We find that persistently overweight identical male twins have approximately $45 - 55$ percent smaller long-term earnings than those who are never overweight. This is driven mainly by a wage effect because persistently overweight men have approximately 30 percent smaller monthly earnings.

Several issues regarding these results should be emphasized. First, our estimated relation sizes are larger than those reported in the previous literature.

TABLE 4 Twin fixed effect regressions, full sample for women and men separately.

	D_Long-term earnings			D_Average employment months			D_Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panel A: Women					
BMI-categories									
D_Persistently overweight (β_1)	-0.298***	-0.290***	-0.270***	-0.780**	-0.785**	-0.769**	-0.170**	-0.162**	-0.148
	(0.106)	(0.105)	(0.104)	(0.318)	(0.319)	(0.319)	(0.082)	(0.082)	(0.081)
D_Previously overweight (β_2)	-0.353	-0.365	-0.378	-0.388	-0.382	-0.349	-0.175	-0.185	-0.204
	(0.221)	(0.221)	(0.216)	(0.478)	(0.480)	(0.484)	(0.148)	(0.149)	(0.142)
D_Subsequently overweight (β_3)	-0.118	-0.109	-0.092	-0.658***	-0.663***	-0.653***	-0.029	-0.021	-0.008
	(0.067)	(0.067)	(0.066)	(0.217)	(0.218)	(0.219)	(0.044)	(0.044)	(0.042)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.020	0.024	0.037	0.007	0.007	0.009	0.149	0.153	0.135
Joint test of controls	-	0.087	0.000	-	0.742	0.329	-	0.012	0.000
$\beta_1 = \beta_2 = \beta_3$	0.219	0.204	0.170	0.744	0.736	0.724	0.128	0.112	0.078
$\beta_1 = \beta_2 + \beta_3$	0.468	0.438	0.387	0.635	0.644	0.680	0.838	0.792	0.691
$\beta_1 = \beta_2$	0.803	0.733	0.613	0.444	0.435	0.422	0.976	0.889	0.722
$\beta_1 = \beta_3$	0.112	0.111	0.108	0.716	0.717	0.730	0.070	0.068	0.065
$\beta_2 = \beta_3$	0.277	0.240	0.177	0.578	0.565	0.538	0.322	0.269	0.167
Number of observations	1182	1182	1182	1182	1182	1182	1182	1182	1182
				Panel B: Men					
BMI-categories									
D_Persistently overweight (β_1)	-0.021	-0.021	0.039	-0.012	-0.012	0.048	-0.040	-0.040	0.006
	(0.095)	(0.095)	(0.092)	(0.248)	(0.248)	(0.246)	(0.058)	(0.058)	(0.056)
D_Previously overweight (β_2)	-0.311**	-0.307**	-0.251	-0.781	-0.783	-0.761	-0.153	-0.145	-0.094
	(0.148)	(0.146)	(0.140)	(0.436)	(0.435)	(0.443)	(0.107)	(0.104)	(0.096)
D_Subsequently overweight (β_3)	-0.047	-0.046	-0.037	-0.296	-0.297	-0.297	-0.035	-0.032	-0.023
	(0.082)	(0.082)	(0.081)	(0.219)	(0.220)	(0.220)	(0.051)	(0.051)	(0.050)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.185	0.186	0.184	0.176	0.172	0.150	0.534	0.561	0.707
Joint test of controls	-	0.463	0.000	-	0.895	0.139	-	0.042	0.000
$\beta_1 = \beta_2 = \beta_3$	0.128	0.130	0.100	0.171	0.167	0.122	0.512	0.532	0.534
$\beta_1 = \beta_2 + \beta_3$	0.039	0.041	0.035	0.029	0.028	0.024	0.202	0.227	0.241
$\beta_1 = \beta_2$	0.045	0.046	0.033	0.078	0.077	0.067	0.269	0.291	0.275
$\beta_1 = \beta_3$	0.772	0.791	0.407	0.264	0.263	0.175	0.936	0.890	0.630
$\beta_2 = \beta_3$	0.076	0.076	0.136	0.272	0.271	0.302	0.269	0.278	0.469
Number of observations	782	782	782	782	782	782	782	782	782

Notes:Full sample includes both fraternal and identical twins. Dependent variables are long-term earnings (log), average employment months (months), and monthly earnings (log), calculated as the average over 1990-2009. Explanatory variables are persistently overweight, previously overweight, and subsequently overweight (omitted category = never overweight), determined by BMI measured in 1975, 1981, and 1990 with BMI threshold > 25 for being overweight. The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. All dependent, explanatory, and control variables are within-twin differenced. Standard errors are in parentheses (heteroscedasticity robust). **p<0.05, ***p<0.01.

However, it is useful to note that unlike many previous papers, we have been careful not to add control variables that themselves may be affected by individuals' overweight history. The reason for this is that we are interested in captur-

TABLE 5 Twin fixed effect regressions, identical twin sample for women and men separately.

	D_Long-term earnings			D_Average employment months			D_Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Women				
BMI-categories									
D_Persistently overweight (β_1)	-0.397**	-0.391**	-0.401**	-1.159**	-1.120**	-1.107**	-0.296	-0.297	-0.309
	(0.185)	(0.189)	(0.178)	(0.545)	(0.553)	(0.553)	(0.218)	(0.216)	(0.214)
D_Previously overweight (β_2)	-0.318	-0.311	-0.259	-0.841	-0.798	-0.701	-0.196	-0.197	-0.174
	(0.269)	(0.272)	(0.288)	(0.994)	(1.004)	(1.031)	(0.102)	(0.100)	(0.105)
D_Subsequently overweight (β_3)	-0.080	-0.078	-0.085	-0.792**	-0.774**	-0.783**	-0.059	-0.060	-0.064
	(0.118)	(0.117)	(0.113)	(0.371)	(0.369)	(0.363)	(0.085)	(0.085)	(0.084)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.187	0.216	0.163	0.067	0.081	0.079	0.258	0.238	0.398
Joint test of controls	-	0.526	0.002	-	0.222	0.118	-	0.853	0.007
$\beta_1 = \beta_2 = \beta_3$	0.291	0.309	0.248	0.795	0.814	0.808	0.238	0.226	0.304
$\beta_1 = \beta_2 + \beta_3$	0.997	0.994	0.852	0.645	0.663	0.722	0.742	0.747	0.579
$\beta_1 = \beta_2$	0.761	0.760	0.605	0.731	0.729	0.671	0.496	0.497	0.370
$\beta_1 = \beta_3$	0.121	0.130	0.095	0.520	0.548	0.566	0.178	0.175	0.155
$\beta_2 = \beta_3$	0.379	0.391	0.539	0.960	0.981	0.935	0.110	0.104	0.222
Number of observations	435	435	435	435	435	435	435	435	435
					Panel B: Men				
BMI-categories									
D_Persistently overweight (β_1)	-0.552***	-0.550***	-0.459***	-0.537	-0.533	-0.371	-0.333***	-0.331***	-0.292***
	(0.192)	(0.192)	(0.174)	(0.455)	(0.457)	(0.443)	(0.103)	(0.105)	(0.103)
D_Previously overweight (β_2)	-0.228	-0.223	-0.212	-0.566	-0.559	-0.536	-0.074	-0.071	-0.060
	(0.137)	(0.139)	(0.143)	(0.345)	(0.350)	(0.350)	(0.083)	(0.086)	(0.088)
D_Subsequently overweight (β_3)	-0.193	-0.194	-0.185	-0.352	-0.354	-0.339	-0.079	-0.080	-0.078
	(0.152)	(0.152)	(0.151)	(0.402)	(0.402)	(0.404)	(0.091)	(0.092)	(0.093)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.039	0.042	0.072	0.410	0.432	0.492	0.014	0.017	0.041
Joint test of controls	-	0.403	0.015	-	0.692	0.319	-	0.446	0.095
$\beta_1 = \beta_2 =_{\beta\ 3}$	0.062	0.065	0.171	0.843	0.854	0.885	0.021	0.024	0.059
$\beta_1 = \beta_2 + \beta_3$	0.475	0.475	0.751	0.428	0.433	0.300	0.156	0.159	0.242
$\beta_1 = \beta_2$	0.050	0.051	0.135	0.945	0.952	0.693	0.010	0.011	0.026
$\beta_1 = \beta_3$	0.031	0.034	0.089	0.635	0.645	0.934	0.032	0.036	0.074
$\beta_2 = \beta_3$	0.829	0.860	0.874	0.604	0.620	0.636	0.966	0.935	0.872
Number of observations	267	267	267	267	267	267	267	267	267

Notes: Dependent variables are long-term earnings (log), average employment months (months), and monthly earnings (log), calculated as the average over 1990-2009. Explanatory variables are persistently overweight, previously overweight, and subsequently overweight (omitted category = never overweight), determined by BMI measured in 1975, 1981, and 1990 with BMI threshold > 25 for being overweight. The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. All dependent, explanatory, and control variables are within-twin differenced. Standard errors are in parentheses (heteroscedasticity robust). **$p < 0.05$, ***$p < 0.01$.

ing the total effect of having a history of being overweight on lifetime earnings. This requires that we allow the regression coefficients of the overweight variables to also capture the indirect effects, such as labor supply decisions, occupational choices, and health behaviors (Han et al. 2011).

Second, many of the previous papers have used a short-term measure, such as the hourly rate of pay, as an outcome variable. Comparing our estimates to those is not straightforward. Our long-term outcome measures have a smaller measurement error than the short-term measures. Our outcome measures are also likely to reflect potential consequences that previous measures were unable to capture, such as the wage effects of repeated unemployment spells and switching jobs often.

Finally, our results show that fully controlling for both genetics and family environment is important for men but not for women. This finding provides an explanation for the lack of consensus in the previous literature on men's earnings penalty for being overweight (Cawley 2015). We acknowledge that this finding is not entirely conclusive because the size of the sample of identical twins is relatively small. However, the finding may be useful for studies using genetic risk scores as instruments for obesity (Böckerman et al. 2018), as it suggests that the power of genetic instruments as well as the exclusion restrictions that their validity require may be gender-specific too.

5.5 Robustness

In this section, we report a number of robustness tests. The tables summarizing the results are included in Appendix 5.A.

5.5.1 Alternative BMI-thresholds

So far we have used the standard BMI threshold for being overweight ($BMI \geq 25.00$) when constructing our main explanatory variables. This threshold is assumed to be the same for both sexes and for all survey years. Below, we consider an alternative way of defining the threshold to see whether the threshold specification affects our results.

Our alternative definition for BMI thresholds is the same as the definition used in the main analysis in Chen (2012). We divide BMI distribution into 3-year and gender-specific terciles and use the lower bounds of the highest terciles as the thresholds. The resulting thresholds for women are 21.55 for 1975, 22.35 for 1981, and 24.90 for 1990. For men, the thresholds are 23.41, 24.49, and 25.68 for 1975, 1981, and 1990, respectively. This means that for women the thresholds for 1975 and 1981 are much lower than the standard threshold used in our main estimations. The thresholds are also lower than the standard for men in 1975 and 1981 but not as low as that of women.

Before we describe our results it is useful to briefly discuss the mechanism behind how these decreases in the thresholds could affect the results. First, a lower threshold adds thinner individuals to the overweight history category, thus changing the composition of each overweight history category. What is important is that the change in the threshold(s) somewhat compromises our definition of being persistently overweight. If our results change and become weaker, this does not automatically imply that our baseline findings were fragile. Instead, it may just mean that keeping the threshold at a higher level also during the earlier measurement years is important for measuring the adverse labor market effects of being persistently overweight.

Our results show that using a lower BMI threshold for the definition of overweight matters. When we replicate the OLS estimations of Subsection 5.4.1. and the fixed effects estimations of Subsections 5.4.2. and 5.4.3, we find that our results become less robust across the models. For example, when the tercile threshold specification is used we no longer find a consistent negative association between persistently being overweight and long-term labor market outcomes for women. For men, using the tercile approach also yields a relatively mixed set of results and weakens the findings for the persistently overweight history category.

These results are in line with the previous literature that suggests that it may be difficult to pin down exactly the point at which weight gains start to matter for labor market outcomes (see, for example, Gregory and Ruhm 2011). Within the context of our study, the standard BMI threshold for being overweight seems to capture the longer-term consequences of being overweight relatively well.

5.5.2 Alternative characterization of weight history

As a robustness check, we also consider an alternative way to characterize our main measure of overweight history, namely, the class of individuals who are persistently overweight. In our main analysis, a persistently overweight individual is defined as an individual who has been overweight either in survey year 1975 or 1981 and was overweight in 1990 as well. As a robustness check, we use a stricter definition for being persistently overweight. Here, we call an individual persistently overweight if the individual is overweight in every survey year: 1975, 1981 and 1990. To keep the estimation sample unchanged, we correspondingly modify how we define the class of previously overweight individuals. An individual is called previously overweight if the individual was overweight in either 1975 or 1981 but not in 1990. In this robustness check, the BMI threshold is kept at 25.

Before reporting the results it is again useful to briefly discuss how changing the definition might affect the results. Changing the definition of being persistent overweight changes the composition of two groups. The group of persistently overweight individuals becomes smaller, and the group of previously overweight individuals becomes larger because some from the group of persistently overweight individuals are now transferred to the group of subsequently overweight individuals. If having some overweight history matters for those who were overweight in 1990, we expect our results for the persistently overweight individuals to weaken and for the subsequently overweight individuals to become stronger.

It is also possible that our baseline results for the persistently overweight individuals are mostly driven by those who were consistently overweight over the fifteen-year period from 1975 to 1990 during which the weight measurements took place. If this is the case, we should expect our results for the persistently overweight individuals to become stronger when the definition of persistently overweight is changed as described.

Our results show that characterizing weight history differently changes the results somewhat but in an expected fashion. The results for the subsequently overweight women become statistically stronger, whereas the results for the persistently overweight women become statistically weaker, especially in models with twin fixed effects.

For men, we find relations only when genetic differences are fully controlled for. As for women, the results for the subsequently overweight men become statistically stronger. A key finding from this robustness analysis is that the documented adverse effects of being overweight on long-term labor market outcomes are related to having *some* overweight history in addition to being overweight in 1990.

5.5.3 Alternative outcome measures

So far, we have calculated long-term earnings as the natural logarithm of average annual earnings. As another robustness check, we re-estimate our models using alternative measures for long-term earnings and monthly earnings. The alternative approach we consider here is to calculate the average of the natural logarithm of annual earnings. Taking the logarithm before averaging can increase the number of missing observations because the logarithm is not defined at zero.

We re-estimate our model using the alternative measure of long-term earnings and the corresponding measure of monthly earnings.[22] We find that using these alternative measures slightly reduces the sample size. However, the change in the definition of the outcome variables has a negligible effect on our results.

5.5.4 Additional control for physical condition

Previous literature has suggested that being overweight and physical condition are related, and being physically active may be related to better long-term labor market outcomes (Hyytinen and Lahtonen 2013). To check the robustness of our results, we have also run our estimations with an additional control for individuals' physical condition, measured in 1975. This variable is defined as 1 if the respondent felt short of breath with light physical effort and is zero otherwise.

Our results show that adding a control for physical condition to the regression models does not change our key results reported in Sections 5.4.1- 5.4.3. The results for women are in line with our previous findings, and the results for men are the same as when the data on identical twins are used.

22 For monthly earnings, we have used the natural logarithm of the ratio of the average long-term earnings and the average employment months. Similarly, the alternative approach would be to take the natural logarithm from the nominator and denominator first and then construct the ratio.

5.6 Discussion

Our findings contribute to the existing literature in four ways. First, it is being persistently overweight in early adulthood that results in lower subsequent long-term earnings for women and men. For men, this association is observed only when data on identical twins are used and genetic differences are fully controlled for. Second, we find that the mechanisms underlying the earnings penalties of being persistently overweight are different for women and men. Persistently overweight women have weaker labor market attachment, working annually almost a month less. The corresponding relationship for persistently overweight men is not statistically significant. Instead, persistently overweight men have lower monthly earnings.

Third, our findings indicate that genetic factors affect the association of being overweight and labor market outcomes for men but not for women. For women, the results are relatively robust across all specifications. For men, they are not: persistently overweight men have smaller long-term earnings and lower monthly earnings after family environment and genetics are simultaneously controlled for. This provides an explanation for the lack of consensus in the previous literature regarding the earnings penalty of overweight men (Averett and Korenman 1996; Baum and Ford 2004; Cawley 2004; Sargent and Blanchflower 1994) and, hence, for why the evidence for men has been more mixed (Cawley 2015). It seems that whether and how genetic factors are a source of omitted variable bias differ across women and men. This interpretation is consistent with the view that genetically inherited traits, such as appearance, and other personality traits that are affected by genetics and environment, such as impatience, contribute to weight gains and shape an individual's labor market outcomes.

Our measure for being persistently overweight allows capturing the cumulative effects of being overweight on labor market outcomes. The underlying mechanism may therefore be related to skill formation and/or other permanent unobservable characteristics of individuals: for women, the negative relationship between being persistently overweight and long-term earnings is related to their weaker labor market attachment. For them, the mechanism is likely to be related to something that erodes women's labor market attachment throughout the lifecycle, for example labor supply decisions or fertility. For persistently overweight men, the mechanism seems to be related to something that erodes their earnings power on the labor market but not their labor market attachment throughout their lifecycle. This could reflect an unobservable characteristic, for example high discount rate, or characteristics related to skill formation earlier in life, such as underdeveloped social skills or lower self-esteem.[23] The question whether persistent overweight reduces productivity due to psychological factors (for example weak self-esteem) or physical factors (for example weak health) would be an interesting topic for future research.

[23] It is also possible that weak self-esteem and social and communication skills contribute to overweight.

We find that, unlike men, subsequently overweight women work less than women who are never overweight. This suggests that additional mechanisms may be at work for women. Besides factors that affect their current productivity, one potential mechanism may be discrimination against overweight women and their job applications (Rooth 2009).

Last, we note that because the period of our analysis is long we acknowledge that stereotypes and prejudices evolve and change as the general population become heavier which can also affect the ability to generalize the obtained results across decades.

5.7 Conclusions

We have studied how an individual's weight history, which we characterize by classes of being persistently, previously, and subsequently overweight, affect individual's long-term earnings, average employment months, and monthly earnings using data on a large number of Finnish twins. We find that (i) being persistently overweight in early adulthood results in lower subsequent long-term earnings for both women and men and (ii) that the mechanism generating this negative relation is gender-specific.

Understanding these linkages is important for policy design. When being overweight reduces people's lifetime earnings, as compared to it only reducing current earnings, the implications for lifetime consumption possibilities and for health policy are profound. Furthermore, interventions and policies that aim at reducing the number of people currently overweight or their contemporary discrimination are likely to have heterogeneous effects because the earnings penalty seems to be specifically associated with persistently being overweight.

5.A Appendix

TABLE A1 Summary statistics for the full, fraternal, and identical twin samples.

	Women			Men		
Sample:	Full	Fraternal	Identical	Full	Fraternal	Identical
Panel A: Dependent variables (1990-2009)						
Long-term earnings (log)	9.37	9.36	9.38	9.70	9.72	9.69
Average employment months (months)	7.90	7.91	7.89	8.19	8.23	8.13
Monthly earnings (log)	7.42	7.42	7.43	7.71	7.71	7.70
Panel B: BMI-category dummies						
Never overweight	0.75	0.74	0.76	0.56	0.52	0.63
Persistently overweight	0.10	0.11	0.09	0.24	0.26	0.19
Previously overweight	0.02	0.02	0.02	0.03	0.03	0.03
Subsequently overweight	0.13	0.13	0.13	0.17	0.19	0.15
Panel C: Control variables						
Age in 1990 (years)	38.70	38.61	38.85	39.18	39.20	39.12
Height in 1975 (cm)	163.48	163.65	163.20	176.55	176.71	176.24
Schooling (years)	12.21	12.18	12.26	12.24	12.12	12.46
Number of diseases 1975	0.50	0.49	0.52	0.39	0.38	0.42
Number of observations	2364	1494	870	1564	1030	534
Number of twin pairs	1182	747	435	782	515	267

Note.Variables in Panel A in addition to age and schooling variables in Panel C are from FLEED. Varibles in Panel B in addition to the height, number of diseases, and employment information in 1975 are from the Older Finnish Twin Cohort Study.

TABLE A2 Within-pair variation in data (means of absolute differences).

	Women			Men		
Sample:	Complete	Fraternal	Identical	Complete	Fraternal	Identical
Dependent variables (1990-2009)						
Long-term earnings (log)	0.68	0.74	0.57	0.65	0.72	0.52
Average employment months (months)	2.28	2.39	2.09	1.89	2.08	1.51
Monthly earnings (log)	0.44	0.47	0.38	0.42	0.44	0.38
BMI-category dummies						
Never overweight	0.24	0.28	0.18	0.32	0.38	0.21
Persistently overweight	0.13	0.15	0.09	0.25	0.28	0.19
Previously overweight	0.04	0.05	0.03	0.06	0.06	0.06
Subsequently overweight	0.20	0.23	0.16	0.26	0.29	0.22
Control variables						
Height in 1975 (cm)	3.23	4.22	1.51	3.71	4.67	1.86
Schooling (years)	1.54	1.75	1.18	1.68	1.85	1.34
Number of diseases 1975	0.57	0.59	0.53	0.52	0.56	0.46
Number of twin pairs	1182	747	435	782	515	267

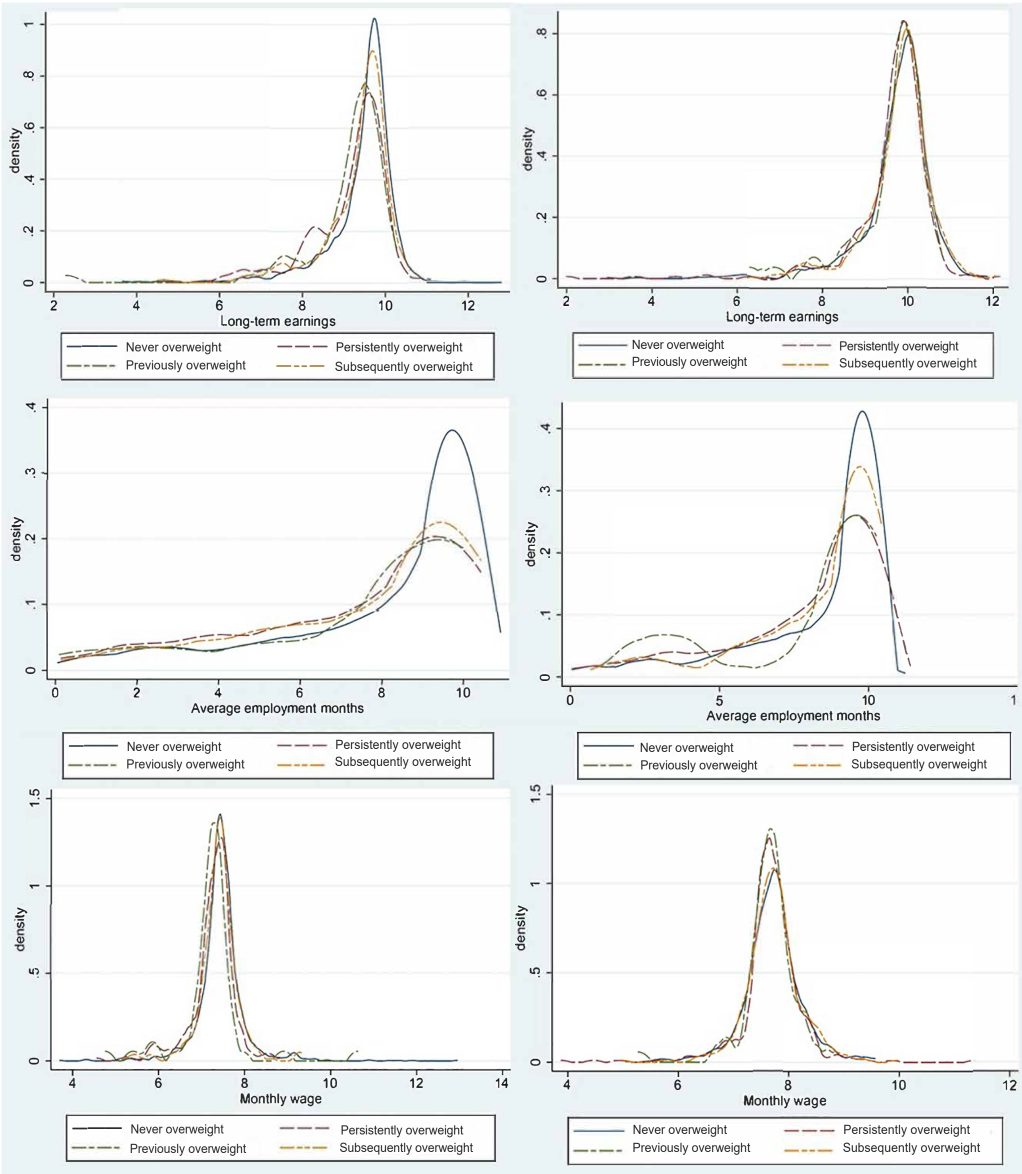

FIGURE A1 Kernel density estimates of long-term earnings (log), average employment months (months), and monthly wages (log) conditional on different obesity history categories for women (left panel) and men (right panel).

TABLE A3 Means in different overweight categories

	Women					Men				
	BMI-category					BMI-category				
	Never overweight	Persistently overweight	Subsequently overweight	Previously overweight	p-value of F-test	Never overweight	Persistently overweight	Subsequently overweight	Previously overweight	p-value of F-test
Panel A: full sample										
BMI										
BMI 1975	20.1	25.4	21.8	24.4	< 0.01	21.2	25.3	22.4	24.3	< 0.01
BMI 1981	20.6	28.0	22.9	25.0	< 0.01	21.8	26.9	23.5	25.4	< 0.01
BMI 1990	21.6	32.0	26.9	23.3	< 0.01	22.6	28.2	26.4	24.2	< 0.01
Age in 1990	38.5	39.8	39.1	39.2	< 0.01	38.7	40.6	38.7	39.9	< 0.01
Dependent variables (1990-2009)										
Long-term earnings (log)	9.42	9.12	9.30	9.06	< 0.01	9.72	9.63	9.77	9.60	< 0.01
Average employment months (months)	8.02	7.40	7.63	7.73	< 0.01	8.31	7.89	8.29	7.90	< 0.01
Monthly earnings (log)	7.46	7.26	7.39	7.22	< 0.01	7.71	7.68	7.73	7.65	< 0.01
Number of observations	1766	241	306	51		874	369	272	49	
Panel B: sample of fraternal twins										
BMI										
BMI 1975	20.1	25.3	21.9	24.1	< 0.01	21.3	25.4	22.4	24.3	< 0.01
BMI 1981	20.6	28.0	24.9	24.9	< 0.01	21.9	26.9	23.5	25.4	< 0.01
BMI 1990	21.7	30.3	27.0	23.2	< 0.01	22.7	28.2	26.3	24.2	< 0.01
Age in 1990	38.4	39.6	38.9	39.5	< 0.01	38.7	40.5	38.6	40.2	< 0.01
Dependent variables (1990-2009)										
Long-term earnings (log)	9.42	9.10	9.33	8.99	< 0.01	9.73	9.71	9.74	9.42	< 0.01
Average employment months (months)	8.01	7.39	7.81	7.67	< 0.01	8.36	8.03	8.29	7.29	< 0.01
Monthly earnings (log)	7.45	7.27	7.41	7.20	< 0.01	7.71	7.73	7.58	7.58	< 0.01
Number of observations	1102	160	197	35		540	265	192	33	
Panel C: sample of identical twins										
BMI										
BMI 1975	20.0	25.8	21.5	24.9	< 0.01	21.1	25.2	22.3	24.3	< 0.01
BMI 1981	20.4	28.1	22.7	25.1	< 0.01	21.7	26.8	23.7	25.5	< 0.01
BMI 1990	21.4	30.3	26.8	23.5	< 0.01	22.6	28.3	26.5	24.3	< 0.01
Age in 1990	38.6	40.2	39.6	38.5	< 0.01	38.6	40.8	39.0	39.3	< 0.01
Dependent variables (1990-2009)										
Long-term earnings (log)	9.43	9.16	9.25	9.22	< 0.01	9.72	9.43	9.84	9.67	< 0.01
Average employment months (months)	8.04	7.40	7.32	7.87	< 0.01	8.22	7.53	8.29	9.18	< 0.01
Monthly earnings (log)	7.46	7.26	7.37	7.34	< 0.01	7.72	7.56	7.81	7.78	< 0.01
Number of observations	664	81	109	16		334	109	80	16	

TABLE A4 Main analysis - OLS regressions, identical twin sample for women and men separately.

	Long-term earnings			Average employment months			Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Women				
BMI-categories									
Persistently overweight (β_1)	-0.268**	-0.256**	-0.146	-0.638**	-0.529	-0.375	-0.198***	-0.207***	-0.121
	(0.116)	(0.113)	(0.105)	(0.303)	(0.293)	(0.279)	(0.068)	(0.071)	(0.067)
Previously overweight (β_2)	-0.210	-0.219	-0.097	-0.167	-0.244	-0.086	-0.191***	-0.185***	-0.085
	(0.171)	(0.170)	(0.192)	(0.743)	(0.742)	(0.750)	(0.053)	(0.055)	(0.064)
Subsequently overweight (β_3)	-0.175**	-0.160**	-0.076	-0.722**	-0.619**	-0.524	-0.095	-0.100	-0.028
	(0.074)	(0.075)	(0.075)	(0.281)	(0.288)	(0.286)	(0.052)	(0.055)	(0.050)
Controls									
Age	N	Y	Y	N	Y	Y	N	Y	Y
Height	N	Y	Y	N	Y	Y	N	Y	Y
Education	N	N	Y	N	N	Y	N	N	Y
Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.014	0.021	0.415	0.019	0.070	0.196	0.000	0.001	0.212
Joint test of controls	-	0.499	0.000	-	0.006	0.000	-	0.260	0.000
$\beta_1 = \beta_2 = \beta_3$	0.778	0.761	0.850	0.777	0.889	0.834	0.271	0.322	0.455
$\beta_1 = \beta_2 + \beta_3$	0.578	0.556	0.906	0.756	0.679	0.772	0.334	0.410	0.930
$\beta_1 = \beta_2$	0.766	0.849	0.813	0.534	0.704	0.702	0.927	0.777	0.678
$\beta_1 = \beta_3$	0.480	0.467	0.569	0.830	0.816	0.692	0.189	0.176	0.224
$\beta_2 = \beta_3$	0.847	0.746	0.916	0.478	0.632	0.577	0.150	0.237	0.463
Number of observations	887	887	887	887	887	887	887	887	887
					Panel B: Men				
BMI-categories									
Persistently overweight (β_1)	-0.292**	-0.315**	-0.171	-0.697**	-0.543	-0.315	-0.162**	-0.199***	-0.112
	(0.131)	(0.135)	(0.128)	(0.348)	(0.332)	(0.319)	(0.070)	(0.071)	(0.067)
Previously overweight (β_2)	0.255**	0.228	0.276***	0.961**	0.760	0.759	0.060	0.072	0.131
	(0.113)	(0.124)	(0.106)	(0.438)	(0.434)	(0.419)	(0.076)	(0.082)	(0.074)
Subsequently overweight (β_3)	0.121	0.087	0.023	0.066	0.046	-0.025	0.084	0.060	0.009
	(0.117)	(0.119)	(0.102)	(0.330)	(0.338)	(0.318)	(0.067)	(0.067)	(0.056)
Controls									
Age	N	Y	Y	N	Y	Y	N	Y	Y
Height	N	Y	Y	N	Y	Y	N	Y	Y
Education	N	N	Y	N	N	Y	N	N	Y
Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.002	0.005	0.015	0.012	0.062	0.169	0.007	0.005	0.073
Joint test of controls	-	0.001	0.000	-	0.000	0.000	-	0.005	0.000
$\beta_1 = \beta_2 = \beta_3$	0.001	0.002	0.010	0.004	0.026	0.081	0.002	0.002	0.032
$\beta_1 = \beta_2 + \beta_3$	0.000	0.001	0.005	0.004	0.021	0.064	0.003	0.004	0.015
$\beta_1 = \beta_2$	0.000	0.001	0.003	0.001	0.008	0.025	0.012	0.007	0.010
$\beta_1 = \beta_3$	0.003	0.006	0.133	0.052	0.132	0.439	0.001	0.001	0.086
$\beta_2 = \beta_3$	0.330	0.348	0.050	0.073	0.151	0.101	0.797	0.907	0.154
Number of observations	548	548	548	548	548	548	548	548	548

Notes: Dependent variables are lifetime earnings (log), average employment months (months), and monthly earnings (log), calculated as the average over 1990-2009. Explanatory variables are persistently overweight, previously overweight, and subsequently overweight (omitted category = never overweight), determined by BMI measured in 1975, 1981, and 1990 with BMI threshold > 25 for being overweight. The control variables are age and age squared in 1990, height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. Standard errors are in parentheses (clustered by twin pair). **$p < 0.05$, ***$p < 0.01$.

TABLE A5 Alternative thresholds - Within-twin regressions, identical twin sample.

	D_Long-term earnings			D_Average employment months			D_Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Women				
BMI-categories									
D_Persistently overweight (β_1)	-0.024	-0.022	0.000	-0.263	-0.252	-0.212	-0.067	-0.067	-0.058
	(0.102)	(0.102)	(0.105)	(0.391)	(0.391)	(0.390)	(0.088)	(0.088)	(0.089)
D_Previously overweight (β_2)	0.161	0.169	0.155	0.306	0.349	0.321	0.036	0.036	0.031
	(0.111)	(0.111)	(0.110)	(0.393)	(0.399)	(0.399)	(0.067)	(0.068)	(0.067)
D_Subsequently overweight (β_3)	-0.265	-0.261	-0.232	-1.218**	-1.198**	-1.138**	0.047	0.048	0.058
	(0.141)	(0.140)	(0.139)	(0.526)	(0.522)	(0.527)	(0.166)	(0.168)	(0.166)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.046	0.042	0.092	0.044	0.040	0.067	0.678	0.687	0.775
Joint test of controls	-	0.371	0.003	-	0.154	0.124	-	0.930	0.011
$\beta_1 = \beta_2 = \beta_3$	0.020	0.018	0.041	0.021	0.019	0.032	0.468	0.478	0.576
$\beta_1 = \beta_2 + \beta_3$	0.673	0.708	0.679	0.307	0.347	0.345	0.486	0.492	0.498
$\beta_1 = \beta_2$	0.097	0.087	0.163	0.142	0.124	0.177	0.227	0.234	0.298
$\beta_1 = \beta_3$	0.148	0.148	0.158	0.072	0.072	0.081	0.602	0.603	0.598
$\beta 2 = \beta 3$	0.008	0.008	0.015	0.006	0.005	0.009	0.949	0.950	0.882
Number of observations	435	435	435	435	435	435	435	435	435
					Panel B: Men				
BMI-categories									
D_Persistently overweight (β_1)	-0.408**	-0.407**	-0.318	-0.887	-0.885	-0.750	-0.128	-0.127	-0.085
	(0.197)	(0.197)	(0.172)	(0.451)	(0.452)	(0.435)	(0.108)	(0.109)	(0.106)
D_Previously overweight (β_2)	-0.273**	-0.276**	-0.217**	-0.599	-0.604	-0.513	-0.072	-0.075	-0.045
	(0.117)	(0.117)	(0.106)	(0.367)	(0.367)	(0.370)	(0.084)	(0.085)	(0.080)
D_Subsequently overweight (β_3)	0.016	0.018	0.054	0.038	0.041	0.094	0.053	0.054	0.069
	(0.248)	(0.247)	(0.252)	(0.532)	(0.533)	(0.545)	(0.182)	(0.181)	(0.185)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.104	0.101	0.171	0.148	0.147	0.230	0.649	0.652	0.834
Joint test of controls	-	0.337	0.011	-	0.651	0.404	-	0.398	0.032
$\beta_1 = \beta_2 = \beta_3$	0.337	0.338	0.445	0.204	0.206	0.271	0.668	0.674	0.762
$\beta_1 = \beta_2 + \beta_3$	0.590	0.597	0.592	0.578	0.584	0.580	0.612	0.619	0.615
$\beta_1 = \beta_2$	0.350	0.365	0.456	0.415	0.429	0.497	0.575	0.603	0.682
$\beta_1 = \beta_3$	0.145	0.144	0.207	0.077	0.077	0.109	0.386	0.384	0.467
$\beta_2 = \beta_3$	0.250	0.238	0.292	0.220	0.216	0.257	0.512	0.497	0.554
Number of observations	267	267	267	267	267	267	267	267	267

Notes: Dependent variables are long-term earnings (log), average employment months (months), and monthly earnings (log), calculated as the average over 1990-2009. Explanatory variables are persistently overweight, previously overweight, and subsequently overweight (omitted category = never overweight). The threshold for overweight is defined by dividing BMI distribution into three-year and gender-specific terciles and use the lower bounds of the highest terciles as the thresholds. The resulting thresholds for women are 21.55 for 1975, 22.35 for 1981, and 24.90 for 1990. For men, the thresholds are 23.41, 24.49, and 25.68 for 1975, 1981, and 1990, respectively. The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. All dependent, explanatory, and control variables are within-twin differenced. Standard errors are in parentheses (heteroscedasticity robust). $^{**}p < 0.05$, $^{***}p < 0.01$.

TABLE A6 Alternative characterization of weight history - Within-twin regressions, identical twin sample.

	D_Long-term earnings			D_Average employment months			D_Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Women				
BMI-categories									
D_Persistently overweight (β_1)	-0.362	-0.362	-0.351	-2.011**	-2.013**	-2.011**	-0.065	-0.065	-0.056
	(0.240)	(0.239)	(0.232)	(0.827)	(0.816)	(0.811)	(0.146)	(0.147)	(0.145)
D_Previously overweight (β_2)	-0.283	-0.276	-0.221	-0.914	-0.875	-0.783	-0.143	-0.144	-0.118
	(0.269)	(0.271)	(0.292)	(0.994)	(1.001)	(1.037)	(0.088)	(0.087)	(0.098)
D_Subsequently overweight (β_3)	-0.151	-0.146	-0.156	-0.767**	-0.738**	-0.740**	-0.136	-0.137	-0.145
	(0.103)	(0.104)	(0.102)	(0.347)	(0.348)	(0.347)	(0.116)	(0.115)	(0.115)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Condition	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.353	0.370	0.353	0.033	0.035	0.033	0.385	0.368	0.510
Joint test of controls	-	0.468	0.003	-	0.191	0.107	-	0.901	0.008
$\beta_1 = \beta_2 = \beta_3$	0.596	0.580	0.653	0.306	0.278	0.268	0.816	0.814	0.789
$\beta_1 = \beta_2 + \beta_3$	0.816	0.846	0.935	0.778	0.732	0.685	0.158	0.152	0.188
$\beta_1 = \beta_2$	0.790	0.773	0.672	0.310	0.291	0.269	0.538	0.535	0.644
$\beta_1 = \beta_3$	0.333	0.318	0.356	0.129	0.115	0.113	0.578	0.575	0.491
$\beta_2 = \beta_3$	0.604	0.609	0.814	0.877	0.885	0.966	0.942	0.940	0.803
Number of observations	435	435	435	435	435	435	435	435	435
					Panel B: Men				
BMI-categories									
D_Persistently overweight (β_1)	-0.517	-0.514	-0.410	-0.328	-0.325	-0.143	-0.232**	-0.231	-0.187
	(0.271)	(0.270)	(0.229)	(0.588)	(0.591)	(0.552)	(0.118)	(0.120)	(0.120)
D_Previously overweight (ß2)	-0.192	-0.187	-0.180	-0.522	-0.515	-0.505	-0.039	-0.035	-0.026
	(0.127)	(0.128)	(0.135)	(0.351)	(0.357)	(0.360)	(0.071)	(0.074)	(0.078)
D_Subsequently overweight (ß3)	-0.316**	-0.316**	-0.280**	-0.448	-0.447	-0.385	-0.179**	-0.179**	-0.163**
	(0.141)	(0.141)	(0.137)	(0.380)	(0.381)	(0.382)	(0.080)	(0.081)	(0.080)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Condition	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health and D_Condition	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
	0.140	0.142	0.188	0.404	0.419	0.400	0.060	0.065	0.131
Joint test of controls	-	0.366	0.009	-	0.676	0.310	-	0.403	0.037
	0.066	0.287	0.470	0.928	0.931	0.748	0.093	0.098	0.191
$\beta_1 = \beta_2 + \beta_3$	0.553	0.955	0.799	0.199	0.202	0.130	0.910	0.898	0.988
$\beta_1 = \beta_2$	0.078	0.137	0.231	0.699	0.705	0.449	0.074	0.076	0.156
$\beta_1 = \beta_3$	0.028	0.361	0.497	0.801	0.797	0.593	0.654	0.667	0.843
$\beta_2 = \beta_3$	0.656	0.331	0.482	0.849	0.864	0.767	0.098	0.097	0.130
Number of observations	267	267	267	267	267	267	267	267	267

Note: Dependent variables are long-term earnings (log), average employment months (months), and monthly earnings (log), calculated as the average over 1990-2009. Explanatory variables are persistently overweight, previously overweight, and subsequently overweight (omitted category = never overweight), determined by BMI measured in 1975, 1981, and 1990 with BMI threshold > 25 for being overweight. The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. All dependent, explanatory, and control variables are within-twin differenced. Standard errors are in parentheses (heteroscedasticity robust). **p < 0.05, ***p < 0.01.

TABLE A7 Alternative outcome measures - Within-twin regressions, identical twin sample.

	D_Long-term earnings			D_Average employment months			D_Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panel A: Women					
BMI-categories									
D_Persistently overweight (β_1)	-0.501***	-0.497**	-0.505***	-1.159**	-1.120**	-1.107**	-0.338	-0.343	-0.354**
	(0.192)	(0.195)	(0.187)	(0.545)	(0.553)	(0.553)	(0.186)	(0.187)	(0.180)
D_Previously overweight (β_2)	-0.280	-0.276	-0.228	-0.841	-0.798	-0.701	-0.165	-0.172	-0.140
	(0.238)	(0.240)	(0.250)	(0.994)	(1.004)	(1.031)	(0.100)	(0.099)	(0.103)
D_Subsequently overweight (β_3)	-0.026	-0.024	-0.030	-0.792**	-0.774**	-0.783**	0.053	0.050	0.045
	(0.099)	(0.099)	(0.095)	(0.371)	(0.369)	(0.363)	(0.077)	(0.076)	(0.074)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.079	0.090	0.062	0.067	0.082	0.079	0.180	0.172	0.154
Joint test of controls	-	0.643	0.013	-	0.222	0.118	-	0.359	0.003
$\beta_1 = \beta_2 = \beta_3$	0.069	0.075	0.048	0.795	0.814	0.808	0.089	0.084	0.072
$\beta_1 = \beta_2 + \beta_3$	0.463	0.459	0.360	0.645	0.663	0.722	0.155	0.152	0.097
$\beta_1 = \beta_2$	0.363	0.363	0.265	0.731	0.729	0.671	0.228	0.223	0.140
$\beta_1 = \beta_3$	0.021	0.023	0.014	0.520	0.548	0.566	0.039	0.038	0.025
$\beta_2 = \beta_3$	0.297	0.305	0.431	0.960	0.981	0.935	0.052	0.047	0.090
Number of observations	435	435	435	435	435	435	435	435	435
				Panel B: Men					
BMI-categories									
D_Persistently overweight (β_1)	-0.480***	-0.477***	-0.403***	-0.537	-0.533	-0.371	-0.333***	-0.332***	-0.299***
	(0.142)	(0.143)	(0.130)	(0.455)	(0.457)	(0.443)	(0.099)	(0.100)	(0.102)
D_Previously overweight (β_2)	-0.189	-0.182	-0.179	-0.566	-0.559	-0.536	-0.147**	-0.144**	-0.146
	(0.105)	(0.106)	(0.106)	(0.345)	(0.350)	(0.350)	(0.072)	(0.073)	(0.075)
D_Subsequently overweight (β_3)	-0.064	-0.065	-0.056	-0.352	-0.354	-0.339	-0.040	-0.040	-0.035
	(0.121)	(0.121)	(0.118)	(0.402)	(0.402)	(0.404)	(0.095)	(0.095)	(0.096)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.007	0.008	0.016	0.410	0.432	0.492	0.010	0.012	0.033
Joint test of controls	-	0.222	0.023	-	0.692	0.319	-	0.498	0.160
$\beta_1 = \beta_2 = \beta_3$	0.010	0.012	0.032	0.843	0.854	0.885	0.044	0.046	0.106
$\beta_1 = \beta_2 + \beta_3$	0.150	0.151	0.288	0.428	0.433	0.300	0.264	0.264	0.382
$\beta_1 = \beta_2$	0.024	0.025	0.068	0.945	0.952	0.693	0.034	0.035	0.093
$\beta_1 = \beta_3$	0.004	0.004	0.011	0.635	0.645	0.934	0.021	0.023	0.043
$\beta_2 = \beta_3$	0.338	0.375	0.345	0.604	0.620	0.636	0.305	0.326	0.301
Number of observations	267	267	267	267	267	267	267	267	267

Notes: Dependent variables are long-term earnings (log), average employment months (months), and monthly earnings (log), calculated as the average of log earnings over 1990-2009. Explanatory variables are persistently overweight, previously overweight, and subsequently overweight (omitted category = never overweight), determined by BMI measured in 1975, 1981, and 1990 with BMI threshold > 25 for being overweight. The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), and the number of diagnosed diseases in 1975. All dependent, explanatory, and control variables are within-twin differenced. Standard errors are in parentheses (heteroscedasticity robust). **p < 0.05, ***p < 0.01.

TABLE A8 Additional control for physical condition - Within-twin regressions, identical twin sample.

	D_Long-term earnings			D_Average employment months			D_Monthly earnings		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Women				
BMI-categories									
D_Persistently overweight (β_1)	-0.397**	-0.387**	-0.398**	-1.159**	-1.079	-1.069	-0.296	-0.296	-0.309
	(0.185)	(0.187)	(0.177)	(0.545)	(0.551)	(0.552)	(0.218)	(0.213)	(0.211)
D_Previously overweight (β_2)	-0.318	-0.317	-0.264	-0.841	-0.854	-0.760	-0.196	-0.198	-0.175
	(0.269)	(0.270)	(0.286)	(0.994)	(0.992)	(1.017)	(0.102)	(0.104)	(0.108)
D_Subsequently overweight (β_3)	-0.080	-0.089	-0.093	-0.792**	-0.881**	-0.883**	-0.059	-0.062	-0.065
	(0.118)	(0.115)	(0.113)	(0.371)	(0.372)	(0.367)	(0.085)	(0.090)	(0.089)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.187	0.209	0.160	0.067	0.059	0.058	0.258	0.249	0.396
Joint test of controls	-	0.498	0.002	-	0.174	0.113	-	0.867	0.007
$\beta_1 = \beta_2 = \beta_3$	0.291	0.332	0.268	0.795	0.929	0.917	0.238	0.206	0.270
$\beta_1 = \beta_2 + \beta_3$	0.997	0.949	0.891	0.645	0.524	0.586	0.742	0.749	0.556
$\beta_1 = \beta_2$	0.761	0.788	0.622	0.731	0.808	0.746	0.496	0.485	0.350
$\beta_1 = \beta_3$	0.121	0.143	0.105	0.520	0.730	0.742	0.178	0.154	0.132
$\beta_2 = \beta_3$	0.379	0.399	0.545	0.960	0.978	0.902	0.110	0.101	0.219
Number of observations	434	434	434	434	434	434	434	434	434
					Panel B: Men				
BMI-categories									
D_Persistently overweight (β_1)	-0.552***	-0.549***	-0.459***	-0.537	-0.532	-0.369	-0.333***	-0.331***	-0.291***
	(0.192)	(0.193)	(0.175)	(0.455)	(0.454)	(0.440)	(0.103)	(0.106)	(0.104)
D_Previously overweight (β_2)	-0.228	-0.231	-0.219	-0.566	-0.580	-0.556	-0.074	-0.077	-0.066
	(0.137)	(0.138)	(0.142)	(0.345)	(0.362)	(0.360)	(0.083)	(0.085)	(0.087)
D_Subsequently overweight (β_3)	-0.193	-0.201	-0.192	-0.352	-0.372	-0.358	-0.079	-0.085	-0.084
	(0.152)	(0.152)	(0.150)	(0.402)	(0.403)	(0.403)	(0.091)	(0.093)	(0.094)
Controls									
D_Height	N	Y	Y	N	Y	Y	N	Y	Y
D_Education	N	N	Y	N	N	Y	N	N	Y
D_Health	N	N	Y	N	N	Y	N	N	Y
Tests (p-values)									
$\beta_1 = \beta_2 = \beta_3 = 0$	0.039	0.043	0.074	0.410	0.421	0.478	0.014	0.019	0.046
Joint test of controls	-	0.365	0.014	-	0.641	0.318	-	0.404	0.086
$\beta_1 = \beta_2 = \beta_3$	0.062	0.072	0.188	0.843	0.865	0.883	0.021	0.029	0.067
$\beta_1 = \beta_2 + \beta_3$	0.475	0.523	0.806	0.428	0.398	0.276	0.156	0.191	0.280
$\beta_1 = \beta_2$	0.050	0.058	0.147	0.945	0.911	0.668	0.010	0.012	0.028
$\beta_1 = \beta_3$	0.031	0.037	0.097	0.635	0.680	0.977	0.032	0.043	0.085
$\beta_2 = \beta_3$	0.829	0.854	0.872	0.604	0.624	0.643	0.966	0.943	0.873
Number of observations	267	267	267	267	267	267	267	267	267

Notes: Dependent variables are long-term earnings (log), average employment months (months), and monthly earnings (log), calculated as the average over 1990-2009. Explanatory variables are persistently overweight, previously overweight, subsequently overweight (omitted category = never overweight), determined by BMI measured in 1975, 1981, and 1990 with BMI threshold > 25 for being overweight. The control variables are height in 1975 (in centimeters), schooling in 1975 (in years), the number of diagnosed diseases in 1975, and physical condition. All dependent, explanatory, and control variables are within-twin differenced. Standard errors are in parentheses (heteroscedasticity robust). **p < 0.05, ***p < 0.01.

References

Aaronson, D. 1998. Using sibling data to estimate the impact of neighborhoods on children's educational outcomes. The Journal of Human Resources 33 (4), 915–946.

Abarca-Gomez, L. e. a. 2017. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: A pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. The Lancet 390 (10113), 2627–2642.

Abramitzky, R., Boustan, L. P., and Eriksson, K. 2012. Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. American Economic Review 102 (5), 1832–1856.

Altonji, J. G. and Dunn, T. A. 1996. Using siblings to estimate the effect of school quality on wages. The Review of Economics and Statistics 78 (4), 665–671.

Angrist, J. D. and Pischke, J.-S. 2009. Mostly harmless econometrics: An empiricist's companion. Princeton: Princeton University Press.

Averett, S. 2011. Labor market consequences: Employment, wages, disability and absenteeism. In The Oxford Handbook of the Social Science of Obesity. Ed. by Cawley, J. Vol. 1. Oxford Handbooks. Oxford University Press, 531–552.

Averett, S. and Korenman, S. 1996. The economic reality of beauty myth. Journal of Human Resources 31 (2), 304–330.

Baum, C. and Ford, W. 2004. The wage effects of obesity: A longitudinal study. Health Economics 13, 885–899.

Behrman, J. and Rosenzweig, M. 2001. The returns to increasing body weight. PIER Working Paper 01 (052).

Behrman, J. and Rosenzweig, M. 2002. Does increasing women's schooling raise the schooling of the next generation? American Economic Review 92 (1), 323–334.

Behrman, J., Rosenzweig, M., and Taubman, P. 1996. College choice and wages: Estimates using data on female twins. The Review of Economics and Statistics 78 (4), 672–685.

Benjamin, D. et al. 2012. The promises and pitfalls of genoeconomics. Annual Review of Economics 4 (1), 627–662.

Bhattacharya, J. and Bundorf, K. 2009. The incidence of the healthcare costs of obesity. Journal of Health Economics 28 (3), 649–658.

Böckerman, P., Cawley, J., Viinikainen, J., Lehtimäki, T., Rovio, S., Seppälä, I., Pehkonen, J., and Raitakari, O. 2018. The effect of weight on labor market outcomes: An application of genetic instrumental variables. Health Economics.

Brunello, G. and Beatrice, D. 2007. Does body weight affect wages?: Evidence from Europe. Economics & Human Biology 5 (1), 1–19.

Burkhauser, R. V. and Cawley, J. 2008. Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. Journal of Health Economics 27 (2), 519–529.

Caliendo, M. and Lee, W.-S. 2013. Fat chanceObesity and the transition from unemployment to employment. Economics & Human Biology 11 (2), 121–133. ISSN: 1570-677X.

Cawley, J. 2004. The impact of obesity on wages. The Journal of Human Resources 39 (2), 451–474.

Cawley, J. 2015. An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. Journal of Health Economics 43, 244–268.

Cawley, J. and Meyerhoefer, C. 2012. The medical care costs of obesity: An instrumental variables approach. Journal of Health Economics 31 (1), 219–230.

Chen, A. J. 2012. When does weight matter most? Journal of Health Economics 31 (1), 285–295.

Cools, S., Markussen, S., and Strøm, M. 2017. Children and careers: How family size affects parents' labor market outcomes in the long run. Demography 54 (5), 1773–1793.

Falkner, N. H., Neumark-Sztainer, D., Story, M., Jeffery, R. W., Beuhring, T., and Resnick, M. D. 2001. Social, educational, and psychological correlates of weight status in adolescents. Obesity Research 9 (1), 32–42.

Farooqi, I. S. and O'Rahilly, S. 2007. Genetic factors in human obesity. Obesity Reviews 8, 37–40.

Florin, T. A., Shults, J., and Stettler, N. 2011. Perception of overweight is associated with poor academic performance in US adolescents. Journal of School Health 81 (11), 663–670.

Gregory, C. A. and Ruhm, C. J. 2011. Where does the wage penalty bite? In Economic Aspects of Obesity. University of Chicago Press, 315–347.

Griliches, Z. 1979. Sibling models and data in economics: Beginnings of a survey. Journal of Political Economy 87 (5), S37–S64.

Gwozdz, W., Sousa-Poza, A., Reisch, L. A., Bammann, K., Eiben, G., Kourides, Y., Kovacs, E., Lauria, F., Konstabel, K., Santaliestra-Pasias, A. M., Vyncke, K., and Pigeot, I. 2015. Peer effects on obesity in a sample of European children. IZA Discussion Papers 9051. Institute for the Study of Labor (IZA).

Han, E., Norton, E. C., and Powell, L. M. 2011. Direct and indirect effects of body weight on adult wages. Economics & Human Biology 9 (4), 381–392.

Han, E., Norton, E. C., and Stearns, S. C. 2009. Weight and wages: fat versus lean paychecks. Health Economics 18 (5) (May 2009), 535–548.

Heckman, J. J. 2006. Skill formation and the economics of investing in disadvantaged children. Science 312 (5782), 1900–1902.

Heckman, J. J. and Mosso, S. 2014. The Economics of Human Development and Social Mobility. Annual Review of Economics 6 (1), 689–733.

Hyytinen, A., Ilmakunnas, P., and Toivanen, O. 2013. The return-to-entrepreneurship puzzle. Labour Economics 20, 57–67.

Hyytinen, A. and Lahtonen, J. 2013. The effect of physical activity on long-term income. Social Science & Medicine 96, 129–137.

Jäntti, M. and Lindahl, L. 2012. On the variability of income within and across generations. Economics Letters 117 (1), 165–167.

Jenq, C., Pan, J., and Theseira, W. 2015. Beauty, weight, and skin color in charitable giving. Journal of Economic Behavior & Organization 119, 234–253.

Johansson, E., Böckerman, P., Kiiskinen, U., and Heliövaara, M. 2009. Obesity and labour market success in Finland: The difference between having a high BMI and being fat. Economics & Human Biology 7 (1), 36–45.

Kaprio, J., Artimo, M., Sarna, S., and Rantasalo, I. 1979. The Finnish twin registry: Baseline characteristics. Section I: Materials methods, representativeness and results for variables special to twin studies. Kansanterveystieteen julkaisuja, M (47).

Kline, B. and Tobias, J. L. 2008. The wages of BMI: Bayesian analysis of a skewed treatment-response model with nonparametric endogeneity. Journal of Applied Econometrics 23 (6), 767–793.

Kristjansson, L. A., Sigfusdottir, I. D., and Allegrante, J. P. 2010. Health behavior and academic achievement among adolescents: The relative contribution of dietary habits, physical activity, Body Mass Index, and self-esteem. Health Education & Behavior 37 (1), 51–64.

Lindeboom, M., Lundborg, P., and Klaauw, B. van der 2010. Assessing the impact of obesity on labor market outcomes. Economics & Human Biology 8 (3), 309–319.

Livshits, G., Malkin, I., Williams, F. M. K., Hart, D. J., Hakim, A., and Spector, T. D. 2012. Longitudinal study of variation in body mass index in middle-aged UK females. AGE 34 (5), 1285–1294.

Mobius, M. M. and Rosenblat, T. S. 2006. Why beauty matters. The American Economic Review 96 (1), 222–235.

Morris, S. 2006. Body mass index and occupational attainment. Journal of Health Economics 25 (2), 347–364.

Morris, S. 2007. The impact of obesity on employment. Labour Economics 14 (3), 413–433.

Norton, E. and Han, E. 2008. Genetic information, obesity, and labor market outcomes. Health Economics 17, 1089–1104.

Paraponaris, A., Saliba, B., and Ventelou, B. 2005. Obesity, weight status and employability: Empirical evidence from a French national survey. Economics & Human Biology 3 (2), 241–258.

Pinkston, J. C. 2017. The dynamic effects of obesity on the wages of young workers. Economics & Human Biology 27, 154–166.

Price, J. and Swigert, J. 2012. Within-family variation in obesity. Economics & Human Biology 10 (4), 333–339.

Puhl, R. and Brownell, K. D. 2001. Bias, discrimination, and obesity. Obesity Research 9 (12), 788–805.

Rooth, D.-O. 2009. Obesity, attractiveness, and differential treatment in hiring: A field experiment. Journal of Human Resources 44 (3), 710–735.

Rothman, K. 2008. BMI-related errors in the measurement of obesity. International Journal of Obesity 32, 56–59.

Sabia, J. J. and Rees, D. I. 2012. Body weight and wages: Evidence from Add Health. Economics & Human Biology 10 (1), 14–19.

Sacerdote, B. 2007. How large are the effects from changes in family environment? A study of Korean American adoptees. The Quarterly Journal of Economics 122 (1), 119–157.

Sargent, J. and Blanchflower, D. 1994. Obesity and stature in adolescence and earnings in young adulthood: Analysis of a British birth cohort. Archives of Pediatrics & Adolescent Medicine 148 (7), 681–687.

Segal, N. L. and Allison, D. B. 2002. Twins and virtual twins: Bases of relative body weight revisited. International Journal of Obesity and Related Disorders 26 (4), 437–441.

Silventoinen, K., Jelenkovic, A., Sund, R., Honda, C., Aaltonen, S., Yokoyama, Y., Tarnoki, A. D., Tarnoki, D. L., Ning, F., Ji, F., and al., et 2015. The CODATwins Project: The Cohort Description of Collaborative Project of Development of Anthropometrical Measures in Twins to Study Macro-Environmental Variation in Genetic and Environmental Effects on Anthropometric Traits. Twin Research and Human Genetics 18 (4), 348–360.

Zwijnenburg, P. J., Meijers-Heijboer, H., and Boomsma, D. I. 2010. Identical but not the same: The value of discordant monozygotic twins in genetic research. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 153B (6).

YHTEENVETO (SUMMARY IN FINNISH)

Taloustieteellisiä tutkimuksia terveydenhuoltomarkkinoista

Väitöskirjassa tarkastellaan terveydenhuoltomarkkinoiden erityispiirteiden ja järjestelmässä olevien ominaisuuksien ja kannustinten vaikutuksia palveluntarjoajien hinnoitteluun, tuotettujen terveyspalveluiden laatuun, markkinarakenteseen ja yhteiskunnalliseen hyvinvointiin. Väitöskirja koostuu johdantoluvusta sekä neljästä tutkimuksesta. Johdantoluku käsittelee terveydenhuoltomarkkinoiden erityispiirteitä, esittelee väitöskirjassa käsiteltävät tutkimuskysymykset, teoreettisen viitekehyksen, käytetyn tutkimusaineiston sekä kokoaa väitöskirjan keskeiset tulokset. Väitöskirjan kolme ensimmäistä tutkimusta ovat soveltavia teoreettisia. Kaksi ensimmäistä tutkimusta tarkastelee hinta- ja laatukilpailun vaikutuksia ja tehokkutta epätäydellisesti kilpailluilla markkinoilla, jolla tarjottavat palvelut tai tuotteet ovat vertikaalisesti differoituja ja joilla on sekä julkisia että yksityisiä palveluntarjoajia. Kolmannessa tutkimuksessa tarkastellaan terveydenhuollon korvausjärjestelmien ominaisuuksien ja markkinoillepääsyn sääntelyn merkitystä. Neljäs tutkimus on empiirinen ja siinä tarkastellaan ylipainon keston ja pitkän aikavälin työmarkkinatulemien välistä yhteyttä. Väitöskirjan tutkimuksista toinen ja kolmas ovat väittelijän yksin kirjoittamia. Tutkimuksista ensimmäinen ja viimeinen on kirjoitettu yhteistyössä kahden eri kirjoittajan kanssa.

Ensimmäinen tutkimus tarkastelee laatu- ja hintakilpailun vaikutuksia ja tehokkuutta epätäydellisesti kilpailluilla markkinoilla, joilla tarjottavat palvelut (tuotteet) ovat vertikaalisesti differoituja ja joilla on sekä julkisia että yksityisiä palveluntarjoajia. Tutkimuksen tulokset osoittavat, että tällaisilla sekamarkkinoilla kilpailun vaikutukset palveluiden hintoihin ja toimijoiden laatuun ovat hyvin erilaiset verrattuna tilanteeseen, jossa tarjonta tapahtuisi puhtaasti yksityisten palveluntarjoajien toimesta. Tutkimuksen päätulos on, että tiettyjen kuluttajien preferenssijakauman muotoon, julkisyrityksen tavoitefunktioon sekä tuotantoteknologiaan liittyvien erityisehtojen vallitessa yritysten laadut voivat markkinatasapainossa olla yhteiskunnan näkökulmasta optimaalisesti valittuja. Koska nämä erityisehdot ovat kuitenkin todellisuudessa harvinaisia, on syytä epäillä, että markkinatasapainolaadut ovat useimmiten tehottomia. Yllättävää on, että optimiin tai optimin lähelle voidaan päästä sääntelemättömillä epätäydellisti kilpailluilla sekamarkkinoilla.

Väitöskirjan toinen tutkimus jatkaa ensimmäisen tutkimuksen aihepiirissä ja olettaa, että palvelun tai tuotteen laatu koostuu useasta ominaisuudesta. Toisen tutkimuksen päätulos on, että ensimmäisessä tutkimuksessa esitetyt erityisehdot eivät ole riittäviä tehokkuuden saavuttamiseksi, kun palvelun laatu on moniuloitteinen. Molemmat tutkimukset viittaavat, että julkisyrityksen toiminta oikeassa laatusegmentissä on yhteiskunnallisen hyvinvoinnin näkökulmasta oleellista. Lisäksi molempien tutkimuksen tulokset havainnollistavat, mitkä preferenssijakauman ja palvelun laaduntuotannon tuotantoteknologian ominaisuudet ovat ajaneet aiemman tutkimuskirjallisuuden tuloksia.

Kolmannessa tutkimuksessa tarkastellaan terveydenhuollon korvausjärjestelmien ominaisuuksien ja markkinoillepääsyn sääntelyn merkitystä. Tutkimuksessa rakennetaan malli, jossa palveluntarjoajien tuotteet ovat laadullisesti differoituja ja jossa osa potilaista havaitsee palveluiden laadun epätäydellisesti. Mallissa palveluntarjoajat kilpailevat potilaista laatuvalinnoillaan. Tutkimuksen päätulokset viittaavat, että markkinoillepääsyn sääntely esimerkiksi toimilupien avulla yhdistettynä tietynlaiseen etukäteen määriteltyihin korvauksiin perustuvaan korvausjärjestelmään voi johtaa yhteiskunnan kannalta optimaaliseen tulemaan. Tulokset viittaavat, että vallitsevasta taloudellisesta ympäristöstä riippuen tällainen sääntelymalli voi olla vapaan markkinoilletulon ja monimutkaisemman korvausjärjestelmän yhdistävää sääntelymallia suositeltavampi sääntelymuoto. Lisäksi tulokset korostavat potilaiden havaitseman ja kokeman terveydenhuollon tarjoajien laadututiedon merkitystä, sillä laadun havainnoinnilla on yhteys sääntelyinstrumenttien ja kilpailun vaikuttavuuteen. Tulosten mukaan potilaille tarjottavalla laatuinformaatiolla voi olla erilaisia suoria ja epäsuoria vaikutuksia, ja vaikutusten suunta riippuu siitä, yli- vai alireagoivatko potilaat terveydenhuollon laatuun.

Neljäs tutkimus keskittyy ylipainoisuuden keston ja pitkän aikavälin ansiotulojen ja työmarkkinoille kiinnittymisen välisen yhteyden tutkimiseen. Tulosten mukaan pysyvä ylipaino on yhteydessä alempiin pitkän aikavälin ansiotuloihin sekä naisilla että miehillä. Alempien ansiotulojen taustalla oleva mekanismi on erilainen naisilla ja miehillä. Naisilla alhaisemmat pitkän aikavälin ansiotulot liittyvät heikompaan työmarkkinoihin kiinnittymiseen läpi elinkaaren. Miehillä alhaisemmat pitkän aikavälin ansiotulot taas liittyvät johonkin sellaisiin tekijöihin, joita tutkimusaineistosta ei voida havaita ja jotka heikentävät heidän ansaintakykyään läpi yksilön elinkaaren.