From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2012

Vouchers, Equality and Competition

Schonger, Martin

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01cn69m416s

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Vouchers, equality and competition

Martin Schonger

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Economics
Adviser: Stephen Morris

June 2012

© Copyright by Martin Schonger, 2012.

All rights reserved.

Abstract

Restricted transfers, or "Money follows people", are a policy instrument that combines public provision of private goods with competition between suppliers. Fee-for-service health care and school vouchers are examples. I find that restricted transfers have two unintended consequences, categorical inequality and competition attenuation, which threatens their promise to deliver the advantages of both government (categorical equality) and the market (competition). Most economists believe that competition (Smith, 1776; Hayek, 1968) drives innovation and thus productivity growth. Friedman (1955) famously argues that restricted transfers allow for higher quality at lower cost than government production.

The intuition behind the two unintended consequences is as follows: the quality a consumer receives depends not only on her direct spending (price paid), but also on her acquisition activities. Examples of acquisition activities are travel, search, information gathering, and bargaining. Holding price constant, the more a consumer is willing to engage in acquisition activities, the higher quality she will receive. But, as Southworth (1945) points out, restricted transfers distort consumption. The consumer can, and will, partially undo that distortion by engaging in less acquisition activities compared to what she would do, if the direct spending had been the result of her own volition. Categorical inequality (ch. 1) occurs since for a normal good the poorer a consumer is, the more she is distorted, and thus the less she is willing to engage in acquisition, which undoes categorical equality. Competition attenuation (ch. 2) occurs as the consumption distortion implies that consumers are less willing to engage in acquisition activities than under ordinary circumstances. This means that suppliers operate in a market where consumers are less willing to switch suppliers for that requires acquisition activities.

The result of chapter 1, the positive quality-income correlation, relies on quality being a normal good. Quality, unlike quantity, is merely ordinal. Normality is conventionally defined with respect to the demand function, i.e. linear budget sets. So the question arises what normality means when budget sets are not intrinsically linear. Chapter 3 untangles the definition of normality from linear budget sets, which allows chapter 4 to show that normality of a good is invariant to any order-preserving transformation of its dimension. Thus irrespective of any particular scaling of quality, and shape of the budget set, a consumer's preference is either normal in quality or not.

Acknowledgements

I gratefully acknowledge steady advice, guidance, incredible patience and encouragement by Stephen Morris and Cecilia E. Rouse.

For helpful comments and discussions concerning this thesis, or parts thereof, I would like to thank Marco Battaglini, Roland Benabou, Sylvain Chassang, Avinash Dixit, Michael Evers, Jacob Goldfield, Faruk Gul, Eric Maskin, Gernot Müller, Harvey Rosen, Mike Rothschild, Bob Rothschild, Chris Sims, and Matthias Wibral.

I would like to thank participants at the following seminars for helpful comments on different chapters of this thesis: The Microeconomic Theory Seminar, the Industrial Relations Graduate Labor Lunch, the Microeconomic Student Seminar and the Public Finance Working Group, all of Princeton University, and the Economics Research Seminar and the Economics Workshop, both of Lancaster University.

Help in the final write-up of this dissertation from Marlon Litz-Rosenzweig and James Binfield is gratefully acknowledged.

For encouragment and moral support, apart from the above, I would like to thank my family, and Gaspard Curioni and Kyla Henriksen.

All errors are mine.

Für meine Frau Mama

Contents

Abstract iii

Acknowledgements iv

List of Figures ix

1 Equal Pay for Unequal Medicine **1**

1.1 Introduction 2

1.2 Models 5

1.2.1 Inter-provider disparity 5

1.2.2 Intra-provider variation 7

1.3 Conclusion 11

References 16

2 Vouchers attenuate competition **17**

2.1 Introduction 18

2.2 The model 23

2.2.1 Policy instruments and parameters 24

2.2.2 Market values 27

2.2.3 Symmetric and non-zero equilibria 27

2.2.4 Existence 28

2.3 Equivalence of vouchers and price floors 29

2.4 Competition Theorem 30

2.4.1 Equilibrium and the competition theorem 30
2.4.2 Perfect competition (τ=0) 33
2.4.3 Imperfect competition (τ>0) 35
2.5 Conclusion 40
Appendix 46
References 58

3 Normal goods **59**
3.1 Introduction 60
3.2 Two characterizations of normality 61
3.3 Defining normality without convexity 64
3.4 Defining normality on the preference itself 67
3.5 Normality in both goods implies convexity 68
3.5.1 Convexity does not imply that both goods are normal 70
3.6 Conclusion 71
Appendix 72
References 72

4 Transformation of a dimension **75**
4.1 Introduction 77
4.2 Example: Screen area and diagonal 80
4.3 Revealed preference and its representations 83
4.4 Invariance 85
4.4.1 Linear transformations 87
4.4.2 Convex transformations 87
4.4.3 Monotone transformations 89
4.5 Conclusion 90
Appendix 92

References 94

List of Figures

1.1 Nash Bargaining Solution 8

2.1 Unit interval of consumers with a firm at each end. 22

2.2 Proper voucher: Profit function given eq. behavior by $-F$. 38

3.1 Change of the slope of the indifference curve as x_2 increases. 63

3.2 Non-convex preferences 65

3.3 Distance between indifference curves. 67

3.4 Normality implies convexity. 69

3.5 Modified Liebhafsky preferences. 70

4.1 Convex preferences are invariant to convex transformations. 78

4.2 Linear transformation of axis 1. 87

4.3 Monotone transformation of axis 1. 89

Chapter 1

Equal Pay for Unequal Medicine

Abstract: Equal Pay for Unequal Medicine refers to the puzzle of patients receiving different qualities of a medical service even when, due to health coverage, they are charged the same price. The two proximate causes are: across providers quality varies even when prices do not, and, at any given provider patients receive different quality even when paying the same price. Moreover it has been observed that poor or minority patients are more likely to receive lower-quality care. While a positive correlation between income and quality holds in most markets, it is astonishing in this setting as price cannot link income with quality. Previous explanations of the disparities thus resort to additional assumptions like provider dislike of the poor, racism, exogenous geography, cultural misunderstandings or a conveniently signed correlation between income or race and patient preference for quality. This paper explains the fact pattern without such assumptions by the differential willingness of patients to incur acquisition costs. The model shows that even in the absence of patient heterogeneity beyond income, equal health coverage must be expected to lead to unequal treatment.

1.1 Introduction

In its report "Unequal Treatment", the Institute of Medicine (Smedley et al., 2003) reviews a large number of empirical studies and concludes that poor and minority patients receive lower quality medical care than other patients even when paying identical amounts. Identical prices of medical services across hospitals and physicians for large groups of individuals are common due to third-party fee-for-service (FFS) health insurance. For example, most of Medicare is FFS and so regardless of patient income, race or ethnicity, or choice of provider, a specific medical service (i.e. billing code) costs the same[1] Medicare-mandated price. While in ordinary markets, a correlation between quality and socioeconomic indicators is explained by varying willingness to pay for higher quality, in this setting, where price is fixed, one might expect to find no large variations in quality correlated with patients' socioeconomic background. Attempts to attribute quality disparities to a potential correlation between patients' socioeconomic background and their medical condition have not been able to account for more than a fraction of the observed variation (e.g. Jha, 2007).

In addition to variation in quality by income among those who seek care, poor and minority patients seek less care even when on the same insurance plan, which has prompted Richman (2008) to ask whether on net insurance expansions redistribute money from poor to rich. This observation, that the uptake of almost free medical services is positively correlated with income is not formalized in this paper, but could be formalized with a model similar to the ones proposed here and the result would obtain as long as there is at least some small inconvenience cost for taking up the "free" medical service.

Carlisle et al. (1997), using the California Hospital Discharge Data Set, examine 105,000 patients with a diagnosis indicating coronary heart disease. Con-

[1]This assumes, as is the case for Medicare (see Shaviro 2004, p. 15), that patients cannot "top up", that is pay the physician or hospital an extra payment beyond the Medicare-specified fee.

trolling for age, sex, principal diagnostic code, the number of secondary diagnoses, elective, urgent or emergent admission, and insurance type they find that African American and Latino patients are much less likely to undergo three main cardiovascular procedures, namely coronary artery angiography, bypass graft surgery and coronary angioplasty. This discrepancy is manifest for patients whose insurance type is HMO, Medicare, Medicaid or who have no health insurance. Remarkably no such difference can be found comparing patients who are privately insured. The authors point out that cost-sharing under HMOs is typically minimal for these procedures, while private health insurance tends to have more cost sharing. Thus they conclude that the out-of-pocket expense borne by patients are an unlikely explanation for this pattern.

The literature has suggested many explanations for the quality disparity. These include differential physician perception of minorities,[2] residential segregation,[3] racial discordance between provider and patient,[4] different preferences of the poor,[5] low assertiveness, lack of information, the historic legacy of segregated care,[6] provider prejudice,[7] and cultural misunderstandings.[8] The quality disparity can be attributed to two sources: (i) poorer patients frequent worse quality providers, and (ii) at any given provider poorer patients receive less quality. Estimates of how much of the quality disparity can be attributed to either of these two sources are rare. In the case of eye exams for diabetics Baicker et al. (2005) estimate that of the quality disparity found between the treatment black versus white patients receive, 56% is attributable to inter-provider variation, and 44% is due to intra-provider variation. Regardless of

[2]E.g. van Ryn (2000), Abreu (1999), and Joe (1998).

[3]E.g. Chandra (2003), Baicker et al. (2004, 2005) , and Wennberg (2002).

[4]E.g. Sohler et al. (2007). On the other hand, for a study that shows that African Americans receive less treatment regardless of their physician's race see Chen et al. (2001).

[5]E.g. Doescher et al. (2001), and Ayanian et al. (1999).

[6]E.g. Smedley et al. (2003, p.103).

[7]E.g. Chandra and Staiger (2010). Schulman et al. (1999) is an audit study that reveals no concious or at least admitted bias, but does reveal bias in treatment recommendations

[8]E.g. Cooper and Roter (1998).

whether arising from across or within provider quality variation, in a setting where price is governmentally mandated and uniform across providers and patients, price cannot explain systematic quality variation.

This chapter shows that the empirical pattern can be explained without recourse to any assumptions of patient heterogeneity beyond income. Some of the assumptions in the literature, could even be explained endogenously using models in the spirit of those put forward in this chapter. The purpose is not to deny the existence of any of the factors the literature mentions, but rather to follow the scientific tradition of aiming for the simplest explanation possible. Thus the only heterogeneity between patients assumed here is patient income.

The intuition of the chapter is that if there is some variable acquisition cost in addition to the uniform price, then differential patient willingness to incur these acquisition costs will lead to different outcomes even when paying the same. Under the term acquisition cost I subsume travel, search, bargaining, information gathering and similar costs. A fixed price restricts and distorts patients' choice, which they will partially undo by adjusting acquisition. If the restricted good is normal then the willingness to incur acquisition costs given a fixed price is increasing in income.

In the next section this insight is applied and demonstrated in three models that differ in the market structure and supply side. The first model addresses the case of the quality disparity arising from inter-provider variation. It demonstrates how poorer patients will on average end up with lower quality due to their selection of low-quality providers, even when they have the same preferences as their wealthier counterparts. The second and third models address the case of the quality disparity arising from intra-provider variation. Both models explain intra-provider variation by implicit bargaining over quality between patient and provider. Section 3 concludes.

1.2 Models

In the following models I try to explain the empirically observed quality disparities as parsimoniously as possible and thus assume that patients are identical except for income. That assumption is not to be understood as an empirical claim, but rather aims to find out whether quality disparities can be explained solely on the basis of income. Therefore patients have identical preferences and the same medical condition, but there are two levels of income $j \in \{P, W\}$, poor patients with income ω^P and wealthy patients with income ω^W, where $\omega^W > \omega^P > 0$. There are two goods, a medical service with some non-negative scalar quality q, of which each patient consumes one unit and a numeraire good. All patients are on the same fee-for-service health plan, so physicians receive a fixed payment $\underline{p}$ for the medical service or procedure regardless of its quality. Physicians are prohibited from charging the patient in addition to that, and do not charge less as any savings would be kept by the health plan. Patients either have to pay nothing or a co-pay κ, so assume $0 \leq \kappa < \omega^P$. The co-pay is irrelevant for the results and just included so that it is clear that the results hold with or without a copay. Patients have strictly monotone, strictly convex preferences over quality and the numeraire good represented by a twice continuously differentiable utility function u. Their preferences are such that quality is a strictly normal good. Normality is used in the usual sense, that is, if patients could choose expenditure on the medical service, their chosen expenditure would be an increasing function of income. By definition strict normality of good 1 means that for all bundles $u_{12}u_2 - u_{22}u_1 > 0$ holds, where the subscripts denote partial derivatives.

1.2.1 Inter-provider disparity

Consider a simple travel cost model where two health care providers are located on opposite ends of a unit interval. There is a unit mass each of poor patients

and wealthy patients. Both types of patients are uniformly distributed on the unit interval. Patients incur a travel cost τ per distance travelled, so a patient located at i $(0 \leq i \leq 1)$ incurs a travel cost τi to the left firm L, and $\tau(1-i)$ to the right firm R. Provider quality is exogenous and heterogeneous: Firm L offers quality q^L, firm R quality q^R. Without loss of generality assume $q^L < q^R$.

***Proposition* 1.1** (**inter-provider variation**)

Either all patients choose the high-quality provider, or the fraction of poor consumers choosing the low-quality provider is strictly higher than the fraction of rich consumers doing so.

Proof. Denote by $\widetilde{i}^j$ the location of the left-most patient with income ω^j who purchases from firm R. Since firm R offers higher quality for the same co-pay it attracts more than half the patients, i.e. $\widetilde{i}^j \leq \frac{1}{2}$, and therefore it suffices to show that $\widetilde{i}^W = 0$ or $\widetilde{i}^W < \widetilde{i}^P$.

1. $\widetilde{i}^P = 0$: Then $\widetilde{i}^W = 0$, since by normality and the fact that a rich patient must be better off than a poor patient at the same location, $u\left(q^L, \omega^P - \kappa - \tau\widetilde{i}^P\right) \leq u\left(q^R, \omega^P - \kappa - \tau\left(1 - \widetilde{i}^P\right)\right)$ imply $u\left(q^L, \omega^W - \kappa - \tau\widetilde{i}^W\right) < u\left(q^R, \omega^W - \kappa - \tau\left(1 - \widetilde{i}^W\right)\right)$.

2. $\widetilde{i}^P \neq 0$, $\widetilde{i}^W = 0$.

3. $\widetilde{i}^P \neq 0$, $\widetilde{i}^W \neq 0$. Then both $\widetilde{i}^P, \widetilde{i}^W$ are indifferent between purchasing at either firm, i.e. for all j: $u\left(q^L, \omega^j - \kappa - \tau\widetilde{i}^j\right) = u\left(q^R, \omega^j - \kappa - \tau\left(1 - \widetilde{i}^j\right)\right)$. Then either $\widetilde{i}^W < \widetilde{i}^P$, in which case we are done, or $\widetilde{i}^W \geq \widetilde{i}^P$. However, $\widetilde{i}^W$ is strictly better off than $\widetilde{i}^P$, and normality of quality implies that $\omega^W - \kappa - \tau\widetilde{i}^W - \left(\omega^W - \kappa - \tau\left(1 - \widetilde{i}^W\right)\right) > \omega^P - \kappa - \tau\widetilde{i}^P - \left(\omega^P - \kappa - \tau\left(1 - \widetilde{i}^P\right)\right)$, which implies that $\widetilde{i}^W < \widetilde{i}^P$.

□

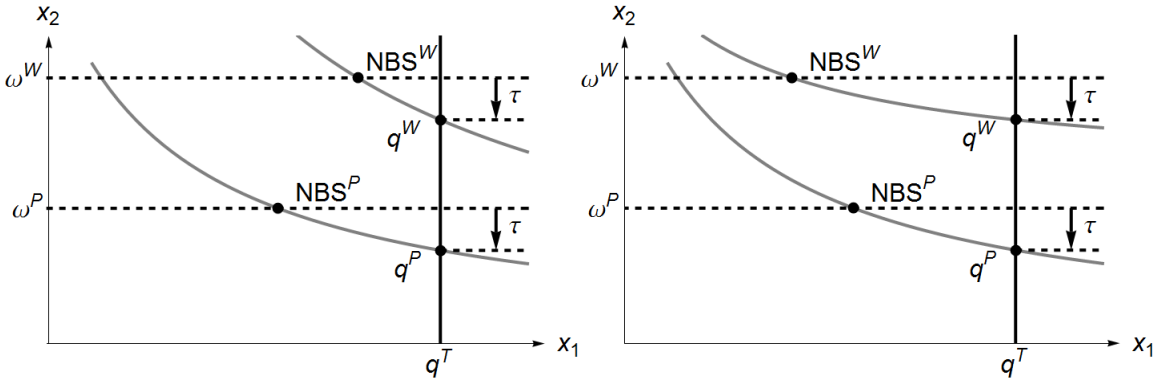
The fraction of wealthy patients $1-\widetilde{i}^W$ choosing the higher quality physician is higher than the fraction $1-\widetilde{i}^P$ of poor patients who do. Another way to look at this is to ask what fraction of the patients in the physician's waiting room are poor. At the low quality physician the fraction of poor patients is $\frac{\widetilde{i}^P}{\widetilde{i}^W+\widetilde{i}^P}$, which is larger than the fraction of poor patients $\frac{\widetilde{i}^P}{2-\widetilde{i}^W-\widetilde{i}^P}$ at the high quality physician.

Note that the self-sorting of patients by income occurs even though there is no correlation between patient income and patient location. If one were to endogenize location choice in this model residential segregation would endogenously emerge with poorer patients living near the low quality physician. Obviously there exists residential segregation for reasons unrelated to health care, the point here is that even in the absence of residential segregation by income the quality disparity should be expected to emerge.

1.2.2 Intra-provider variation

Even more puzzling than the phenomenon that poor patients frequent worse providers than their wealthy counterparts on exactly the same FFS-health coverage, is quality discrimination at a given physician. The physician receives the same payment from all patients, yet wealthy patients receive higher quality than poor patients. This observation can be explained by introducing Nash bargaining between patient and physician. Rather than some haggling over quality in the practice, the bargaining should be imagined as implicit. Given FFS coverage, bargaining is only over quality, not also over price. Physicians have all the bargaining power.[9] The patients' outside option or threat point is to leave the physician's waiting room and seek care at another physician where they know they receive quality q^T for sure. But switching physicians

[9]If patients had all the bargaining power, then there would be no variable acquisition costs and thus all patients would get the same quality. There are no variable acquisition costs in the sense that in the Rubinstein game (1982), which provides a non-cooperative foundation for the Nash-Bargaining solution, the party with all the bargaining power is infinitely patient, that is, it has no costs from waiting.



(a) Normal good. (b) Inferior good

Figure 1.1: Nash Bargaining Solution

is inconvenient and modelled by the switching cost $\tau > 0$. Therefore the patients' threat level of utility is $u\left(q^T, \omega^j - \kappa - \tau\right)$. Physicians observe patient income, this should not be taken literally when applying the model, but rather interpreted to mean that physicians observe some correlates of patient income and use those for statistical discrimination. The physician's threat point is to refuse treatment and to make zero profit. The profit function differs between the two following models that follow.

Intra-provider variation I: provider-effort

In the provider-effort model the physician can vary the amount of effort or time taken to perform the medical service. Quality is increasing in effort, but the effort is costly to the physician. This is captured by the fact that physician profit is decreasing in quality: $\pi(q) = \underline{p} - c(q)$. In order to have a bargaining problem there has to be a potential for bargaining surplus. Thus there must exist at least one type of patient who is strictly better off receiving the quality that makes the physician zero profit rather than the patient's threat point: there exists j such that $u\left(c^{-1}\left(\underline{p}\right), \omega^j - \kappa\right) > u\left(q^T, \omega^j - \kappa - \tau\right)$, where c^{-1} is the inverse of c.

***Proposition* 1.2 (intra-provider variation from provider effort)**

Wealthy patients receive strictly higher quality than poor patients.

Proof. Since the physician has all the bargaining power, in the Nash solution patients receive their threat level utility. Thus they are indifferent between staying or switching to another physician: $u(q, \omega^j - \kappa) = u\left(q^T, \omega^j - \kappa - \tau\right)$ for all j. For strictly monotone and convex preferences this equation implicitly defines q as a function of ω. That function is increasing in income ω if quality is a normal good as figure 1.1 illustrates: in (a) quality is a normal good, while in (b) it is an inferior good. □

Even more surprisingly it is the poorer patient who receives the worse deal. Conventional price discrimination usually means that the poor get the better deal. Here price is administratively fixed by Medicare reimbursement rules, so medical providers can discriminate in quality alone. This quality discrimination has the opposite effect of the usual price discrimination in as much as it gives a worse deal to the poor.

Intra-provider variation II: choice of billing code

In the billing code model the physician cannot vary effort for a given medical service but can choose which of two medical services or procedures to administer. Both procedures are medically justifiable and get reimbursed. Procedure H has a higher quality than procedure L: $q^H > q^L$. The associated billing codes and thus reimbursements differ and are $\underline{p}^L, \underline{p}^H$. The physician incurs cost c^L, c^H for each procedure and his profit is the difference between the reimbursement and the cost. If procedure H is more profitable for the physician then there is no conflict of interest between the physician and the patient as both prefer procedure H which is therefore chosen. Therefore focus on the interesting case where there is a conflict of interest and assume that the profit of procedure L is larger than from procedure H: $\pi(L) = \underline{p}^L - c^L > \pi(H) = \underline{p}^H - c^H \geq 0$. Using the above framework the provider maximizes profit subject to the constraint that the patient receives no less than her threat utility $u\left(q^T, \omega^j - \kappa - \tau\right)$. Assume that it is feasible to do so, i.e. $u\left(q^L, \omega^j - \kappa\right) \geq u\left(q^T, \omega^j - \kappa - \tau\right)$.

***Proposition* 1.3** (intra-provider variation from billing code)
If $u\left(q^L, \omega^W - \kappa - \tau\right) < u\left(q^T, \omega^W - \kappa - \tau\right)$ *and* $u\left(q^L, \omega^P - \kappa - \tau\right) \geq u\left(q^T, \omega^P - \kappa - \tau\right)$ *then wealthy patients receive high quality, poor patients the low quality procedure, else all patients receive the same quality procedure.*

Proof. (i) Suppose that the poorer patient undergoes the higher quality treatment H: Since the provider is foregoing the more profitable treatment L, we can conclude that L would lead to less than the poor patient's threat utility, i.e. $u\left(q^L, \omega^P - \kappa\right) < u\left(q^T, \omega^P - \kappa - \tau\right)$. But then normality of quality implies that $u\left(q^L, \omega^W - \kappa\right) < u\left(q^T, \omega^W - \kappa - \tau\right)$. Therefore if the poor patient receives the less profitable, higher quality treatment H, a fortiori the wealthy patient is also offered H.

(ii) Suppose that the richer patient is offered treatment L and chooses to undergo it: thus her utility under treatment L is at least as large as her threat utility, i.e. $u\left(q^L, \omega^W - \kappa\right) \geq u\left(q^T, \omega^W - \kappa - \tau\right)$. A fortiori by normality $u\left(q^L, \omega^P - \kappa\right) > u\left(q^T, \omega^P - \kappa - \tau\right)$, but then the provider will also only offer q^T to the poor patient who will accept.

(iii) The poorer patient is offered only treatment L, chooses to undergo it, while the wealthy patient is offered and accepts treatment H: The provider is foregoing the more profitable treatment L only for the wealthy patient, thus we conclude that L does not reach the threat utility of the wealthy patient, but does for the poor: $u\left(q^L, \omega^W - \kappa\right) < u\left(q^T, \omega^W - \kappa - \tau\right)$ and $u\left(q^L, \omega^P - \kappa\right) \geq u\left(q^T, \omega^P - \kappa - \tau\right)$. □

A model like this could explain findings like Baicker's (2004) observation that blacks have more money spent on them but receive the less effective treatments.

In both bargaining models in equilibrium poor and rich consumers spend the same on acquisition, that is neither incurs the switching cost τ. Their differential willingness to incur it is what drives the difference.

1.3 Conclusion

This chapter develops a simple explanation for the empirically observed systematic quality disparity between patients of different socioeconomic backgrounds occurring even though price paid is the same for all patients due to fee-for-service (FFS) health coverage. While the empirical pattern of quality inequality would not be surprising in any market, conditional on price we would usually expect the disadvantaged to receive no less quality. The quality disparity arises from two factors, (i) poorer patients frequenting lower quality health care providers,[10] and (ii) poorer patients getting lower quality than their wealthier counterparts who frequent the same provider. To explain (i), inter-provider disparity a travel cost model is developed in this chapter, while (ii), intra-provider disparity is explained using two Nash bargaining models. In all three models what drives the disparity in quality is the disparity in consumer income; consumers were assumed to be identical in all other aspects to show the strength of the theory. As price is fixed by government, the income disparity cannot produce the quality disparity via price. However, as long as quality of the medical service is not an inferior good, the richer a consumer, the higher is her willingness to engage in costly acquisition activities (e.g. search, travel, bargaining). In the travel cost model in equilibrium wealthier patients on average travel further than poor patients do. In the bargaining models in equilibrium no patient switches providers or spends any time on bargaining, but the differential willingness to do so leads wealthier patients to obtain higher quality from the same provider even as they pay the same, possibly zero, co-pay, and the provider gets the same payment from their health insurance.

By proposing this simple explanation for the empirically observed quality disparity and furthermore theoretically predicting that uniform restricted transfers, vouchers or FFS-health coverage will lead to systematic differences in quality received

[10] At the extreme a choice of lower quality could be interpreted as not seeing a health care provider at all, and thus the model can explain lower uptake as for example observed in Baicker (2004).

by income as long as there is some variable acquisition activity this chapter advances the understanding beyond the current state of the literature as for example surveyed in Summers (1989), Bradford and Shaviro (2000), or Currie and Gahvari (2008).

The distinguishing feature of this chapter's explanation of the quality disparity compared to prior explanations is its parsimony: while other explanations need to take recourse to additional assumptions of heterogeneity between patients such as residential location, health status or preferences, this chapter explains the phenomenon by heterogeneity in income alone. This parsimony mirrors Becker (1968) who explains the higher incidence of property crime among the poor without postulating heterogeneous preferences or morals. Furthermore, like Becker this chapter must not be understood to claim that there are no differences between poor and rich other than income, or that such differences could not explain different outcomes. Rather the point is to show that differential outcomes do not necessarily imply the existence and causality of other differences. Future empirical work is needed to distinguish between this parsimonious explanation and other factors on a case-by-case basis.

Similarly, correlations between race and outcomes, rather than income and outcomes can be explained, again like in Becker (1968), by the fact that even nowadays race remains a predictor of permanent income, even when controlling for transitory income. Conventional stories of different mores, genetically or culturally caused health conditions and behaviors, or widespread and shocking levels of racism among health care providers are, while in principle sufficient, again not necessary to explain such differences.

Arrow (1963) famously pointed to information asymmetries in medicine, explaining that understanding of treatment options and their quality is low among patients. If one interprets travel cost metaphorically as information acquisition cost then in this chapter heterogenous and suboptimal levels of information arise endogenously in the travel-cost model, thus endogenizing Arrow's observation.

For policy this chapter is highly relevant as it shows that the conventional wisdom that equalizing access to health care via uniform fee-for-service health coverage should lead to equal quality of treatment received is mistaken. Unequal quality for equal pay is not a puzzle, but follows from standard assumptions of consumer theory. On the contrary equal quality for equal pay for consumers with different incomes would be a puzzle. Such a puzzle could be explained by quality being a constant good,[11] quality not being variable to begin with, the absence of potentially variable acquisition costs or if acquisition requires spending resources that are more costly for the wealthy such as time.

Patients can learn from the observation of inverse quality discrimination that they are likely to get a better deal if they pretend to be rich rather than poor, which would be the opposite recommendation in the standard case of a market with price discrimination. Policymakers may want to rethink their preference for categorical equality. Fee-for-service health coverage in any case is only seemingly equal, but ends up giving more to the wealthy. If categorical equality in the quality of treatment is really desired then FFS would have to be adjusted such that provider reimbursements for each medical service are not uniform, but decreasing in patient income. As an alternative to categorical equality which as this chapter explains is costlier to achieve than one might think, policymakers could consider other options to improve the lives of the poor, as for example Deaton (2002) who calls for policymakers to "relax constraints on poor people tackling low incomes and poor education".

[11] As defined in chapter 3 constant good is a good for which spending is constant in income. For quasi-linear preferences the goods that are not quasi-linear are constant.

References

Abreu, J. M. (1999). Conscious and nonconscious African American stereotypes: impact on first impression and diagnostic ratings by therapists. *Journal of Consulting and Clinical Psychology 67*(3), 387–93.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review 53*(5), 941–73.

Ayanian, J. Z., P. D. Cleary, J. S. Weissman, and A. M. Epstein (1999). The effect of patients' preferences on racial differences in access to renal transplantation. *New England Journal of Medicine 341*(22), 1661–9.

Baicker, K., A. Chandra, and J. S. Skinner (2005). Geographic variation in health care and the problem of measuring racial disparities. *Perspectives in Biology and Medicine 48*(1), S42–53.

Baicker, K., A. Chandra, J. S. Skinner, and J. E. Wennberg (2004). Who you are and where you live: how race and geography affect the treatment of medicare beneficiaries. *Health Affairs Web Exclusive*, 33–44.

Becker, G. S. (1968). Crime and punishment: an economic approach. *Journal of Political Economy 76*(2), 169–217.

Bradford, D. F. and D. N. Shaviro (2000). The economics of vouchers. In C. E. Steuerle (Ed.), *Vouchers and the provision of public services*, Chapter 2, pp. 40–91. Washington, DC: Brookings Institution Press.

Carlisle, D. M., B. D. Leake, and M. F. Shapiro (1997). Racial and ethnic disparities in the use of cardiovascular prcedures: Associations with type of health insurance. *American Journal of Public Health 87*(2), 263–7.

Chandra, A. and J. S. Skinner (2003). Geography and racial health disparities. Working Paper 9513, NBER.

Chandra, A. and D. O. Staiger (2010). Identifying provider prejudice in health care. Working Paper 16382, NBER.

Chen, J., S. S. Rathore, M. J. Radford, Y. Wang, and H. M. Krumholz (2001). Racial differences in the use of cardiac catherization after acute myocardial infarction. *New England Journal of Medicine 344*(19), 1443–9.

Cooper, L. and D. Roter (1998). *Patient-provider communication: the effect of race and ethnicity on process and outcomes of healthcare*, pp. 552–93. Washington, D.C.: National Academies Press.

Currie, J. and F. Gahvari (2008). Transfers in cash and in kind: theory meets the data. *Journal of Economic Literature 46*(2), 333–83.

Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health Affairs 21*(2), 13–30.

Doescher, M. P., B. G. Saver, K. Fiscella, and P. Franks (2001). Racial/Ethnic inequities in continuity of care. *Health Services Research 36*(6), 78–89.

Jha, A. K., D. O. Staiger, F. L. Lucas, and A. Chandra (2007). Do race-specific models explain disparities in treatments after acute myocardial infarction? *American Heart Journal 153*(5), 785–91.

Joe, J. (1998). *The rationing of healthcare and health disparity for the American Indians/Alaska natives*, pp. 528–51. Washington, D.C.: National Academies Press.

Richman, B. D. (2008). Insurance expansions: Do they hurt those they are designed to help? *Health Affairs 26*, 1345–57.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica 50*(1), 97–109.

Schulman, K. A., J. A. Berlin, W. Harless, J. F. Kerner, S. Sistrunk, B. J. Gersh, R. Dube, C. K. Taleghani, J. E. Burke, S. Williams, J. M. Eisenberg, and J. J. Escarce (1999). The effect of race and sex on physicians' recommendations for cardiac catheterization. *New England Journal of Medicine 340*(8), 618–26.

Shaviro, D. (2004). *Who Should Pay for Medicare?* University of Chicago Press.

Smedley, B. D., A. Y. Stith, and A. R. Nelson (2003). Unequal treatment: confronting racial and ethnic disparities in health care.

Sohler, N., L. Fitzpatrick, R. Lindsay, K. Anastos, and C. Cunningham (2007). Does patient-provider racial/ethnic concordance influence ratings of trust in people with HIV infection? *AIDS and Behavior 11*(6), 884–96.

Summers, L. H. (1989). Some simple economics of mandated benefits. *American Economic Review 79*(2), 177–83.

van Ryn, M. and J. Burke (2000). The effect of patient race and socio-economic status on physicians' perceptions of patients. *Social Science and Medicine 50*(6), 813–28.

Wennberg, J. E., E. S. Fisher, and J. S. Skinner (2002). Geography and the debate over Medicare reform. *Health Affairs Web Exclusive*, W96–114.

Chapter 2

Vouchers attenuate competition

Abstract: School vouchers and fee-for-service health insurance combine public funding with competitive production and consumer choice by fully reimbursing the consumer's expenditure up to a fixed amount. For I refer to all such policy instruments as vouchers. A voucher has an income effect and a consumption distortion effect as first observed by Southworth (1945). This chapter combines two insights about vouchers. First, it argues vouchers are de-facto price floors, or more formally they implement the same set of equilibria as a price floor combined with a cash transfer does. Second, the consumption distortion lowers consumers' willingness to engage in costly acquisition activities such as travel, search and bargaining in exchange for higher quality. Combining these insights shows that vouchers attenuate competition. Vouchers that are so low as to merely create an income effect, but no distortion effect, do not attenuate competition. The competition attenuation is shown in a parsimonious model of a market with competing firms and acquisition costs. Vouchers attenuate competition, while quality floors, a hypothetical benchmark that directly targets quality, do not. In the limit, as the acquisition cost parameter goes to zero, competition becomes perfect and remains so even in the presence of a voucher. Thus vouchers attenuate competition if and only if they distort consumption from the consumer's perspective and there are variable acquisition costs. Competition attenuation is deeply worrying from a policy perspective, as competition is often believed to be the main driver of productivity growth, and thereby long-run human welfare.

2.1 Introduction

In most rich and middle-income countries education and health care comprise a large share of government expenditure. For example in the United States more than 36 percent of expenditure at all levels of government is dedicated to health care and education (BEA, 2011). A figure that does not even account for tax expenditures. Given their large share of government expenditures, it is therefore paramount for the fiscal sustainability of both government as a whole, as well as that of the public provision of health care and education in particular, to investigate the drivers or hindrances of productivity in these sectors.

Government production and allocation of private goods such as education and health care promises access for all yet does not allow for competition. Competition, however, is believed by many economists and policymakers to drive innovation and thus productivity. For example Smith (1776) viewed the lack of competition as an enemy of good management, and v. Hayek (1968) even called competition a discovery procedure. Friedman (1955) famously pointed out that it is possible to combine public funding with competition between private producers in his call for universal school vouchers. School vouchers, fee-for-service health care and other restricted transfers are all policy instruments that combine public funding with competition between producers. A hope many have is therefore that such policies will allow for vigorous competition leading to innovation and productivity growth.

Yet Friedman's hope that combining public funding with consumer choice would lead to high quality at low cost does not seem to have worked out in another sector, one where it was actually tried out on a large scale, namely the health care sector. Medicare is a system that, like Friedmanite school vouchers, combines public funding and consumer choice.[1] While Medicare is not universal, it does cover everyone

[1] In health care the unit for which government reimburses is not a year of schooling, as proposed by Friedman for education, but a specific medical procedure as delineated by a billing code.

over 65, and therefore it and its beneficiaries dominate medical demand. In health care this public funding combined with consumer choice seems not to have been conducive to dramatic increases in productivity: As a percentage of GDP spending on health care rose from 4.8% in 1960 to 9.8% in 1985 to 16.5% in 2009 (Congressional Budget Office, 2011). To be sure these numbers in themselves are not conclusive about productivity, an ageing population and increased demand in a richer society have to be take into account, but the increase is so dramatic as to suggest at most low productivity growth. Note that many improvements in mortality and morbidity are not due to health care, but improved living conditions, especially early in life (see Fogel, 2004). Further on the point of the medical cost explosion, Finkelstein (2007) finds that the introduction of Medicare increased spending by over six times as much as small-scale experiments would have predicted. Small scale experiments should capture most of the demand side effect, so the difference observed in the study might be attributed to the supply side.

Baumol and Bowen (1965) famously observe that productivity growth in health care and education is smaller than in other sectors and coined the term cost disease. They suggest to explain it with an almost immutable requirement of labor in such services, pointing out that it is much easier to reduce the amount of labor required in manufacturing than education or health care.

This chapter offers a novel explanation for Baumol's cost disease, that can be an alternative or addition to Baumol and Bowen's on explanation. Like Friedman I consider competition to be a key driver and incentive for improvements in productivity. But this chapter argues that competition in sectors where consumers have the choice among providers, and providers compete only over non-price dimensions, competition might be much weaker than one would expect. The point that vouchers and fee-for-service health coverage weaken competition compared to markets is new to the literature. Friedman (1955), calling for school vouchers, suggested that they

would lead to "more effective competition" and thus a "more efficient utilization of [..] resources". Friedman did not provide a formal model. To be sure his point is that vouchers would allow more competition than government production and distribution[2], while this chapter uses the market as a benchmark. This chapter asks whether in markets that are dominated by consumers using restricted transfers such as school vouchers or fee-for-service health coverage one should expect levels of competition similar to those encountered in full-fledged markets.

Both school vouchers and fee-for-service health coverage give the beneficiary a fixed amount of money that may only be spent on the target good at a supplier of her choice. For brevity henceforth this chapter refers to these policy instruments as vouchers. To fix ideas consider the empirically most relevant case of a discrete good where government wants to ensure high quality. This discrete good could be a year of schooling or a medical service such as a particular surgery. Let quality be measured by a non-negative scalar. Take the quantity of the discrete good each individual consumes as given and equal to one unit. Government wants to increase the quality consumed and does so indirectly via the easily observable and court-verifiable price of the discrete good. If all consumers have a voucher of the same amount, then, since consumers cannot keep any savings, a firm has no reason to charge less than the amount of the voucher. Thus the only way to attract more consumers is for firms to offer higher quality.

Southworth (1945) was the first to observe that any voucher that distorts consumption leads to a lower utility for the consumer compared to an unrestricted cash transfer, . Often this utility loss is referred to as a deadweight loss. This is misleading, whether the utility loss is a deadweight loss or not, can only be judged in general equilibrium. For example if the voucher addresses a positive consumption ex-

[2]Note that even with government production and distribution there is already some very limited form of competition, as local governments compete for residents à la Tiebout (1956).

ternality of the good, then it may reduce an existing, rather than create, a deadweight loss.

A key insight of this paper is that the consumption distortion has another implication apart from the utility loss: The consumption distortion lowers the consumer's willingness to engage in costly acquisition activities such as travel, search, bargaining or information gathering in return for higher quality. This insight and its consequences for competition are formally investigated in this chapter. For parsimony the simplest example of a market model with acquisition costs, a travel cost model, is employed.

Three policy instruments are compared: markets in conjunction with cash transfers, vouchers, and quality floors in conjunction with cash transfers. Under the market firms are free to choose any price and quality they desire. A voucher means that government reimburses consumers up to the value of the voucher. In equilibrium a voucher that is higher than the market price can indeed raise quality above the market level, which confirms the conventional wisdom. A quality floor means that government legally mandates a minimum quality, leaving firms free to choose any quality above the floor.

In the model imperfect competition is caused by travel costs. The higher the travel cost parameter, the more imperfect the competition. Firms earn location rents which in aggregate are just equal to the travel cost parameter. The central result of this chapter, the *competition theorem*, states that firms' location rents are higher under any proper voucher than they are both in the fully-fledged, ordinary market and under any quality floor. Moreover location rents are strictly increasing in the voucher and the effect is of first-order. Put differently, a larger voucher has the same effect on competition as higher travel costs would.

The empirical literature on vouchers is vast, for an excellent survey see Currie and Gahvari (2008). Theoretical investigations of vouchers are much rarer,

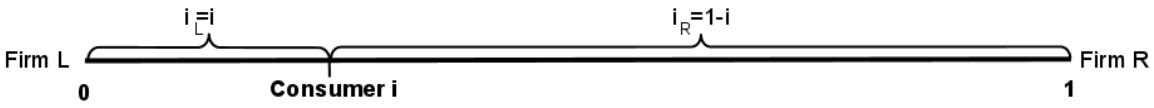


Figure 2.1: Unit interval of consumers with a firm at each end.

and most focus on the interaction of vouchers with other issues such as political economy (Epple and Romano, 1996; Blomquist and Christiansen, 1999), income distribution effects (Epple, Newlon and Romano, 2002), or residential segregation (Nechyba, 1996). Theoretical investigations of vouchers per se start with Southworth, other articles taking a more general focus, besides the already mentioned Currie and Gahvari article, include Summers (1989), and Bradford and Shaviro (2000).

I do not discuss the reasons why public policy should aim to increase consumption of a rival good in the first place. Arguments that have been proposed include positive externalities of consumption, time-inconsistent consumers (Mulligan and Philipson, 2000), the Samaritan's dilemma (Buchanan, 1975), merit goods (Musgrave, 1959), good-specific altruism as a biological fact (Hanson, 2008), political camouflage (Whitmore, 2002), screening for true necessity (Nichols and Zeckhauser, 1982), addressing credit constraints hindering efficient investment in human capital, and channeling resources to a household member other than the household head.

Understanding this competition-attenuating effect is important for policymakers considering employing such mechanisms. If such mechanisms are chosen, then policymakers should understand how to construct them and accompanying market regulations in a way that lessens the competition-weakening effect.

Section 2.2 develops the model, section 2.3 explains that vouchers can be understood as price floors, and section 2.4 states and explains the competition theorem. Section 2.5 concludes.

2.2 The model

A unit mass of consumers is uniformly distributed on the unit interval. A firm is located at each end, firm L at 0, and firm R at 1. Firms produce a single output, a discrete good with a single-dimensional quality q. Individuals buy at most one unit of the discrete good. Each firm F can produce a single quality $q_F \geq 0$ of the good. Firms do not observe a consumer's location i. A firm's production cost for one unit of good 1 is given by a twice continuously differentiable, strictly increasing and weakly convex function c of quality q_F , with $c(0) = 0$. Cost is linear in the quantity produced. The firm charges price p_F for one unit of the good. A firm's strategy is a price-quality pair $s_F = (p_F, q_F)$. Its markup is therefore a function $\mu = \mu(s_F) = p_F - c(q_F)$. Demand is equal to the number of a firm's customers as each customer purchases one unit. Consumers have a choice whether and where to purchase. This choice is a function of the prices and qualities offered by both firms. Thus a firm's demand is an implicit function of both its strategy and its competitor's strategy: $d = d(s_F, s_{-F})$, and so is its profit: $\pi = \pi(s_F, s_{-F}) = d(s_F, s_{-F}) \cdot (p_F - c(q_F))$.

Acquisition costs are modeled as a travel costs: To purchase at firm F consumer i incurs a cost of τi_F, where $i_L \equiv i$ and $i_R \equiv 1 - i$. The acquisition cost parameter τ satisfies $\tau \geq 0$. The acquisition cost parameter is small compared to consumers' income ω.[3] Individuals consume two goods, zero units or one unit of good 1, the discrete good, and good 2, a numeraire good. The numeraire good is traded on a competitive market. Consumers have homogeneous preferences on $\mathbb{R}^2_+$, represented by a twice continuously differentiable utility function u. The first dimension represents the quality of the discrete good, while the second dimension represents the quantity

[3] Assume at the very least that $\tau < \frac{2}{3}\omega$.

of the numeraire good. Preferences are strictly monotone and strictly convex. Good 1 is strictly normal but not superior.[4]

The game structure is as follows: Firms simultaneously choose prices and qualities, then consumers observe these prices and qualities and choose whether and from which firm to buy good 1. If a consumer chooses to buy from firm F her budget constraint is $p_F + x_2 + \tau i_F \leq \omega$. If she does not buy good 1 then her budget constraint is $x_2 \leq \omega$. Consumers maximize utility. If a consumer is indifferent between buying from firm L and R, assume that she chooses the closer one. To avoid corner solutions assume that the marginal utility at zero is infinite, and consumers always prefer to consume both goods over just the numeraire good or over just good 1.[5]

2.2.1 Policy instruments and parameters

Three kinds of government policies are investigated: market, vouchers, price floors and quality floors. To enable a comparison between these policies where government expenditure is kept constant, all policies except vouchers are considered in combination with non-negative cash transfers. A policy can be a legal mandate which reduces the strategy space S of the firms, which in the absence of such mandates is $\mathbb{R}_0^+ \times \mathbb{R}_0^+$.

Market

The market policy means that there is no government intervention except for possibly a cash transfer T. In this case the policy does not reduce the strategy space so it remains $S = \mathbb{R}_0^+ \times \mathbb{R}_0^+$. The consumer's budget constraint when purchasing good 1 from firm F is $p_F + x_2 + \tau i_F \leq \omega + T$.

[4] I distinguish between strongly and strictly, where strictly refers to the analytical analagon on almost all points. Stricty monotone $u_1, u_2 > 0$, strictly convex $2u_1u_2u_{12} - u_{11}u_2^2 - u_{22}u_1^2 > 0$. Good 1 strictly normal $u_{12}u_2 - u_{22}u_1 > 0$, not superior $u_{12}u_1 - u_{11}u_2 > 0$.

[5] For all $q \geq 0, x_2 \geq 0$: $u_2(q,0) = u(0,x_2) = \infty$, $u\,(0,0) > u\,(-,x_2)$, where – denotes no consumption of good 1 at all.

Voucher

Government reimburses each consumer for the amount spent on good 1 up to some pre-announced maximum reimbursement of $V > 0$. When a consumer spends more than the voucher amount on good 1, the literature speaks of "topping up". Topping up is allowed in the model, but changing that assumption does not alter the substance of the results of the paper. To summarize, if a consumer purchases the good from firm F she receives a government reimbursement of $\min\{V, p_F\}$, and her budget constraint is thus $p_F + x_2 + \tau i_F \leq \omega + \min(V, p_F)$. The strategy space of firms remains $S = \mathbb{R}_0^+ \times \mathbb{R}_0^+$.

Price floor

A price floor $\underline{p}$ is a legal mandate that requires firms to sell at a price no lower than $\underline{p}$. The strategy space is therefore $S = \left\{p | p \geq \underline{p}\right\} \times \mathbb{R}_0^+$. Government can combine a price floor with a non-negative cash transfer T. The consumer's budget constraint when buying good 1 is $p_F + x_2 + \tau i_F \leq \omega + T$. To ensure affordability government only considers combinations of price floors and cash transfers which satisfy $\underline{p} < \omega + T - \frac{3}{2}\tau$.

Quality floor

A quality floor $\underline{q}$ is a legal mandate that requires firms to sell quality at a level no lower than $\underline{q}$. The strategy space is therefore $S = \mathbb{R}_0^+ \times \left\{q | q \geq \underline{q}\right\}$. Government can combine a quality floor with a non-negative cash transfer T. Thus the consumer's budget constraint when purchasing from firm F is $p_F + x_2 + \tau i_F \leq \omega + T$. To ensure affordability government only considers combinations of quality floors and cash transfers which satisfy $\omega + T - \frac{3}{2}\tau > c\left(\underline{q}\right)$.

Notational convention

I adapt the notational convention that if there is no voucher then $V = 0$, if there is no price floor then $\underline{p} = 0$, and if there is no quality floor then $\underline{q} = 0$. So regardless of the policy instrument, the consumer's budget constraint can be written as $p_F + x_2 + \tau i_F \leq \omega + T + \min(V, p_F)$. By monotonicity the budget constraint binds, which allows expressing a consumer's numeraire consumption (given that she buys good 1) as a function of her location i, choice of firm F, and its price p_F by $x_2(i, F, p_F) = \omega + T + \min\{V, p_F\} - p_F - \tau i_F$.

Improper - the analogue of non-distortionary in general equilibrium

Southworth's analysis of vouchers is one of partial equilibrium as it is a demand side analysis only. A voucher of V, depending on its size, may or may not be seen by the consumer as equivalent to a cash transfer of the same size. Loosely speaking one can think of the consumer as being constrained to spend V on the target good, with the equivalence between vouchers and cash transfers occurring when this constraint is not binding. Another way to think about it is to observe that the voucher is equivalent to the cash transfer if the consumption is non-distorted from the consumer's perspective. In a general equilibrium model, however, firms may take into account the constraints consumers face and thus the constraints might be binding on firms rather than consumers. Thus I define the analogue of a non-distortionary or non-binding voucher for general equilibrium and call such vouchers *improper*. Moreover, it is useful to have a definition which is applicable to policy instruments other than vouchers:

***Definition* 2.1** (**improper voucher**)

We say that a policy instrument is ***improper,*** *if the set of equilibria implemented by it is a subset of the set of equilibria implemented by the cash transfer which has the same budgetary cost for government as the policy instrument does in equilibrium.*

Basically this means that vouchers that only affect equilibrium via their income effect are called improper, the same logic goes for other policy instruments. For example a quality floor by itself has no income effect, and thus is improper only if it does not change the set of equilibria at all. A quality floor in conjunction with a cash transfer T would be improper, if just giving the cash transfer T would lead to the same equilibria. The above definition of improper is likely suitable for a wide range of models of the market beyond the particular examples presented here.

2.2.2 Market values

To describe and understand the equilibria implemented by different policy instruments and parameters it is useful to have a name for the particular values of price and quality that will turn out to be the unique equilibrium candidates under a market policy with cash transfer T. Call these values the *market values* and denote them by (p°, q°). The market values take the income effect of cash transfers into account, and thus depend only on consumers' post-transfer income I which is determined by the pre-transfer income ω and the policy instrument and parameter. The marginal rate of transformation (MRT) is just the derivative of the cost function.

***Definition* 2.2** (**market values**)
Given post-transfer income I implicitly define the market value of quality $q^\circ = q^\circ(I)$ by $MRS\left(q^\circ, I - \frac{3}{2}\tau - c\left(q^\circ\right)\right) \equiv MRT\left(q^\circ\right)$, and the market value of price p° by $p^\circ \equiv c\left(q^\circ\right) + \tau$.

The appendix (p. 46) shows that $q^\circ(I)$ is indeed well-defined by the above. Note that $p^\circ > 0$ and $q^\circ > 0$.

2.2.3 Symmetric and non-zero equilibria

Call an equilibrium (s_L^*, s_R^*) symmetric if both firms set the same price and quality:

***Definition* 2.3** (**symmetry**)

An equilibrium (s_L^*, s_R^*) *is* ***symmetric*** *if* $s_L^* = s_R^*$.

(s^*, s^*) denotes a symmetric equilibrium. I call an equilibrium non-zero if all consumers buy both goods, and both firms set their price and quality to be larger than zero:

***Definition* 2.4** (**non-zero**)

An equilibrium (s_L^*, s_R^*) *is* ***non-zero*** *if* $d(s_L^*, s_R^*) + d(s_R^*, s_L^*) = 1$, $p_F^* > 0$, $q_F^* > 0$, *and* $x_2(i, L \text{ if } i \leq d(s_L^*, s_R^*) \text{ else } R, p_F^*) > 0$ *for all* F *and* i.

If equilibrium is symmetric it can be written as (p^*, q^*, p^*, q^*).

2.2.4 Existence

These assumptions are not strong enough to ensure existence of equilibrium for all preferences and values of τ compared to ω. The assumption $\tau < \frac{2}{3}\omega$ is particularly weak. That particular upper limit of τ merely ensures that any equilibrium is non-zero. For higher values of τ it could happen that there is an interval of consumers in the middle between the two firms that cannot afford to purchase the target good in equilibrium. But $\tau < \frac{2}{3}\omega$ does not ensure existence. The appendix gives an example of such non-existence of equilibrium. The appendix also conjectures that for any preference, endowment ω and policy subject to the above assumptions, there is a $\overline{\tau}$ small enough such that for all τ smaller than $\overline{\tau}$, equilibrium exists. The results in this chapter characterizing equilibrium rely on necessary conditions for an equilibrium only. But for ease of exposition assume existence of equilibrium:

***Assumption* 2.1**

Assume that preferences and parameters, including policy parameters, are such that equilibrium exists.

2.3 Equivalence of vouchers and price floors

To understand the intuition behind the results of this paper, as well as their robustness beyond the present model, it is most helpful to observe that a voucher is equivalent to the combination of a price floor and cash transfer where both have the same value as the voucher. The games firms play under a voucher versus a price floor differ: with a voucher it is legal for a firm to set a price below the voucher amount, whereas with a price floor it is not. But the voucher and the combination of quality floor and cash transfer give rise to the same set of equilibria, a result I call equivalence:

***Proposition* 2.1** (**equivalence**)
The set of equilibria implemented by a voucher of V is identical to the set of equilibria implemented by price floor of $\underline{p}$ in conjunction with a cash transfer of T, where $\underline{p} = T = V$.

The proof is in the appendix (p. 52). To understand the logic behind the result note that consumers cannot keep any savings from purchasing at a price lower than the voucher and therefore demand remains the same when V instead of p_F is charged: $d(p_F, q_F, p_{-F}, q_{-F}|V) = d(V, q_F, p_{-F}, q_{-F}|V)$. That is under a voucher demand is perfectly inelastic with respect to price as long as price is lower than V. As a consequence strategies s_F with $p_F < V$ are weakly dominated. The game firms play under the price floor and cash transfer combination is essentially the same game as is played under the voucher once weakly dominated strategies are eliminated. As the elimination of weakly dominated strategies is not always innocuous, this observation is just intended to give an understanding of the result. The proof itself relies on the fact that in equilibrium firms' demand is positive, and then any strategy (p_F, q_F) with $p_F < V$ yields strictly lower profit than the strategy (V, q_F).

2.4 Competition Theorem

2.4.1 Equilibrium and the competition theorem

In the model the acquisition cost parameter τ exogenously affects the degree of competition: the larger τ the less intense competition is. In addition to τ of course other factors could have an impact on competition, thus τ itself is not a suitable measure for the intensity of competition in equilibrium, more appropriate and standard measures of competition might be the markup or aggregate rents. The markup can in principle differ between firm L and firm R, but as shown in the appendix (p. 47) any equilibrium is symmetric. The proof mainly relies on the fact that since the game is symmetric imitating the competitor's strategy is always feasible, and thus can never be more profitable than a firm's equilibrium strategy. Thus there is only one equilibrium markup, and thus I use the equilibrium markup to measure competition. Moreover markup is a good measure, as in equilibrium the markup equals aggregate rents, the obvious alternative measure. This is due to the fact that there is a unit mass of consumers, and in equilibrium all consumers buy the target good as shown in the appendix (2.3). Note that a higher markup indicates *weaker* competition, and that a markup of zero indicates perfect competition.

***Theorem* 2.1** (**competition**)

The markup in an equilibrium that is implemented by a cash transfer, an improper voucher, or a quality floor in conjunction with a cash transfer equals τ.

The markup in an equilibrium that is implemented by a proper voucher is strictly larger than τ if and only if there are acquisition costs.

The competition theorem states that a cash transfer by itself leads to an equilibrium markup of τ. An improper voucher, price floor or quality floor by definition implements the same equilibrium as the corresponding cash transfer, and therefore all improper policies have an equilibrium markup of τ. This is true irrespective

of whether acquisition costs are zero or positive. A proper voucher, and its equivalent proper price floor, lead to an equilibrium markup larger than τ. This cannot be due to an income effect, as a cash transfer of the same size still has a markup equal to τ. The cash transfer and the voucher both raise the quality consumed in equilibrium, but per definition a proper voucher raises it more. Under the cash transfer for the consumer in the middle, $i = \frac{1}{2}$, the marginal rate of substitution equals the marginal rate of transformation. Under the proper voucher the marginal rate of substitution is smaller than marginal rate of transformation. Might the latter thus explain why a voucher raises the markup, while a cash transfer of the same size does not? Here quality floors are a useful benchmark. A proper quality floor in conjunction with a cash transfer like a voucher raises quality due to a positive income effect and a distortion forced on the consumer, and thus for proper quality floors, like for proper vouchers, the marginal rate of substitution is smaller than the marginal rate of transformation. But as the competition theorem states the markup under a proper quality floor equals τ. Thus the fact that a voucher increases the markup, and therefore attenuates competition, is not at all related to its income effect, nor solely due to the fact that it distorts consumption. It is due to the fact that it distorts consumption, the fact that it does so indirectly via price rather than quality, and the fact that acquisition costs are not zero. The fact that the competition attenuation only occurs if consumption of quality is distorted by targeting indirectly via price, rather than directly quality itself is reminiscent of results in trade theory and environmental economics that also state that directly targeting the variable of interest is better than doing so indirectly. What logic lies behind this? As argued before, a consumer cannot keep any savings from a price below the voucher, thus demand is perfectly price inelastic at prices below the voucher amount. For a proper voucher the equilibrium price equals the amount of the voucher, since a proper voucher is essentially a proper, i.e. binding, price floor. Thus a proper voucher, being a binding price floor, eliminates competition in price. A firm

that wants to deviate from equilibrium to attract more consumers can do so only by offering higher quality. If the market is frictionless to begin with, that is if there are no acquisition costs, this feature of a voucher is harmless, as for any arbitrarily small increase in quality above the equilibrium level the firm captures the entire market. In other words, for zero acquisition costs, under the market policy demand is perfectly elastic in both price and quality, a voucher makes the downward price elasticity zero at price equal to or below the voucher amount, but this has no effect on overall competition since the demand remains perfectly elastic with respect to quality. The fact that a proper voucher eliminated competition in price becomes harmful only in the presence of variable acquisition costs. Distorting consumption means that under both proper vouchers and proper quality floors, the demand elasticity with respect to price is larger than the demand elasticity with respect to quality normalized by its marginal cost. Thus the quality floor which eliminates competition in quality, the dimension where it is weaker, does not alter competition, while the voucher, which eliminates competition in price, the dimension where it is stronger, does reduce competition.

The table gives an overview over the characteristics of the equilibrium under each policy. The most important observation is in the last column: equilibrium markup $\mu(s^*)$ equals τ, except when there are positive acquisition costs and a proper voucher, then the equilibrium markup is larger than τ.

policy instrument	parameter(s)	equilibrium price p^*	equilibrium quality q^*	consumption distortion? MRS	equilibrium markup $\mu(s^*)$
cash transfer	T	$= p^\circ(\omega + T)$	$= q^\circ(\omega + T)$	$= MRT$	$= \tau$
quality floor					
improper iff	$\underline{q} \leq q^\circ(\omega + T)$	$= p^\circ(\omega + T)$	$= q^\circ(\omega + T)$	$= MRT$	$= \tau$
proper iff	$\underline{q} \leq q^\circ(\omega + T)$	$= c\left(\underline{q}\right) + \tau$	$= \underline{q}$	$< MRT$	$= \tau$
voucher					
improper iff	$V \leq p^\circ(\omega + V)$	$= p^\circ(\omega + V)$	$= q^\circ(\omega + V)$	$= MRT$	$= \tau$
proper iff	$V > p^\circ(\omega + V)$	$= V$	$< q^\circ(\omega + V)$	$< MRT$	$= \tau$ | $> \tau$ (zero | positive acquisition cost)

2.4.2 Perfect competition (τ=0)

Consider the case where there are no acquisition costs ($\tau = 0$), which results in perfect competition. Perfect competition prevails in the sense, that in equilibrium, as this section shows, demand is infinitely elastic with respect to both price and quality and firms' profits are zero. Profits are zero since for $\tau = 0$ the game is a Bertrand game with a two-dimensional strategy space, price and quality, rather than just price as in the standard Bertrand game. That profits are zero is implied by the following lemma, which states that the equilibrium markup is zero:

***Lemma* 2.1** (**zero markup**)

If $\tau = 0$ *then for any equilibrium* (s^*, s^*) *the markup is zero:* $\mu(s^*) = 0$.

For the proof see the appendix (p. 54). Since the markup is zero price equals the cost of production: $p^* = c(q^*)$. Given that fact the equilibrium strategy must maximize consumers' utility. Thus the equilibrium can be found by maximizing consumers' utility subject to $p = c(q)$, in addition to $p \geq \underline{p}$ and $q \geq \underline{q}$. The latter two constraints can be simplified to $q \geq \max\left\{c^{-1}\left(\underline{p}\right), \underline{q}\right\}$. Note that the objective function is quasi-concave in u. The first-order condition for an interior maximum is $MRS(q, \omega + T + V - c(q)) = MRT(q)$. But this is the definition of the market value given a post-transfer income of $I = \omega + T + V$. Thus the unique critical point of the quasi-concave maximization problem is the market value of quality $q = q^\circ(\omega + T + V)$. Thus the market value is the solution of the maximization problem iff it satisfies the constraints, otherwise a corner solution must occur and the respective constraint must bind.

Market

Under the market policy $\underline{p} = \underline{q} = 0$, and thus the solution is interior and given by the market values $s^* = (p^\circ(\omega + T), q^\circ(\omega + T))$. Since in equilibrium $MRS(q^*, \omega + T - p^*) = MRT(q^*)$ holds, consumption is not distorted.

Quality floor

Consider a quality floor of $\underline{q} > 0$ in conjunction with a cash transfer of $T \geq 0$.

If $\underline{q} \leq q^{\circ}(\omega + T)$ then the problem of maximizing consumers' utility does not differ from the problem given the market policy, except for this non-binding constraint. Hence equilibrium is again $s^* = (p^{\circ}(\omega + T), q^{\circ}(\omega + T))$. Since the same unique equilibrium occurs as under the market policy with the same cash transfer, the quality floor is improper.

Otherwise $\underline{q} > q^{\circ}(\omega + T)$, thereby the unique critical point $q^{\circ}(\omega + T)$ is not in the domain of the maximization problem. Therefore a corner solution occurs and thus $q^* = \underline{q}$, and as the markup is zero equilibrium price is $p^* = c(\underline{q})$.

To summarize, a quality floor of $\underline{q}$ is proper if and only if $\underline{q} > q^{\circ}$. Consumption of quality is not distorted for an improper quality floor, but as $MRS(q^*, \omega + T - p^*) < MRT(q^*)$ it is distorted upward for a proper quality floor.

Voucher

Consider a voucher of V. By the equivalence proposition the equilibria can be found by instead examining a price floor of $\underline{p} = V$ in conjunction with a cash transfer of $T = V$.

If $\underline{p} \leq p^{\circ}(\omega + V)$, which is equivalent to $c^{-1}(\underline{p}) \leq q^{\circ}(\omega + T)$, then the maximization problem does not differ from the problem under the market policy, except for this non-binding constraint. Hence equilibrium is $s^* = (p^{\circ}(\omega + V), q^{\circ}(\omega + V))$. Since the same unique equilibrium occurs as under the market policy in conjunction with a cash transfer of $T = V$, the price floor, and ergo the voucher is improper.

Otherwise $\underline{p} > p^{\circ}(\omega + V)$, which is equivalent to $c^{-1}(\underline{p}) > q^{\circ}(\omega + T)$, and thus the unique critical point $q^{\circ}(\omega + T)$ is not in the domain of the maximization problem. Therefore a corner solution occurs and thus $s^* = (\underline{p}, c^{-1}(\underline{p}))$.

To summarize, a voucher of V is proper if and only if $V > p^{\circ}(\omega + V)$. Consumption of quality is not distorted for an improper voucher, but as $MRS(q^*, \omega + V - p^*) < MRT(q^*)$ it is distorted upward for a proper voucher.

Comparison of instruments

If acquisition costs are zero vouchers and quality floors both lead to an equilibrium markup of zero. Both policy instruments lower willingness to incur acquisition costs compared to the market, but this is inconsequential as there are no such costs to be incurred in the first place.

2.4.3 Imperfect competition (τ>0)

The fact that any equilibrium is non-zero and symmetric implies that in equilibrium consumer $i = \frac{1}{2}$ is indifferent between purchasing at either firm. Moreover, since equilibrium is non-zero for sufficiently small deviations by a firm from its equilibrium strategy, an indifferent consumer continues to exist. If an indifferent consumer exists then demand of firm L is given by the location of the indifferent consumer. Thus the demand function of firm L conditional on equilibrium play by firm R is implicitly defined by $H(d, s_L) = 0$, where $H(d, s_L) \equiv u(q_L, x_2(d, L, p_L)) - u(q^*, x_2(d, R, p^*))$. Note that the former holds only if the deviation satisfies $p_L > V$ and $q > \underline{q}$. Since $x_2(d, L, p_L) = \omega + T + V - p_L - \tau i$, and $x_2(d, R, p^*) = \omega + T + V - p^* - \tau(1 - i)$. Then $H(d, p_L, q^*)$ is differentiable with respect to p_L and $H(d, p^*, q_L)$ q_L with respect to q_L. Therefore by the implicit function theorem:

$$\left.\frac{\partial d}{\partial p_L}\right|_{(s_L, s^*)} = -\frac{u_2(q_L, x_2(d, L, p_L))}{\tau(u_2(q_L, x_2(d, L, p_L)) + u_2(q^*, x_2(d, R, p^*)))} \tag{2.1}$$

$$\left.\frac{\partial d}{\partial q_L}\right|_{(s_L, s^*)} = \frac{u_1(q_L, x_2(d, L, p_L))}{\tau(u_2(q_L, x_2(d, L, p_L)) + u_2(q^*, x_2(d, R, p^*)))} \tag{2.2}$$

To evaluate these equations at the equilibrium values, note that, since equilibrium is symmetric and non-zero, $d(s^*, s^*) = \frac{1}{2}$. The derivative of demand with respect to price evaluated in equilibrium is:

$$\left.\frac{\partial d}{\partial p_L}\right|_{(s^*,s^*)} = -\frac{1}{2\tau} \tag{2.3}$$

This means that if firm L marginally raises its price, its demand declines by half of the inverse of the travel cost. The derivative with respect to quality evaluated in equilibrium is:

$$\left.\frac{\partial d}{\partial q_L}\right|_{(s^*,s^*)} = \frac{1}{2\tau} MRS\left(q^*, x_2\left(\frac{1}{2}, L, p^*\right)\right) \tag{2.4}$$

This means that if firm L marginally raises quality its demand increases by the inverse of the travel cost multiplied by the marginal rate of substitution between quality and the numeraire. The partial derivatives of profit evaluated at the equilibrium values are, as long as $p_L^* > V$ and $q_L^* > \underline{q}$:

$$\left.\frac{\partial \pi}{\partial p_L}\right|_{(s^*,s^*)} = \frac{1}{2\tau}\left(\tau - \mu\left(q^*\right)\right), \tag{2.5}$$

$$\left.\frac{\partial \pi}{\partial q_L}\right|_{(s^*,s^*)} = \frac{1}{2}\left(MRS\left(q^*, x_2\left(\frac{1}{2}, L, p^*\right)\right)\frac{\mu\left(q^*\right)}{\tau} - MRT\left(q^*\right)\right), \tag{2.6}$$

Market

Since equilibrium satisfies $p_L^* > 0$ and $q_L^* > 0$, an equilibrium implemented by the market policy can only occur at an interior point of the profit-maximization problem, and thus partial derivatives of profit must be zero:

$$\left.\frac{\partial \pi}{\partial p_L}\right|_{(s^*,s^*)} = 0 \Rightarrow \mu\left(s^*\right) = \tau \tag{2.7}$$

$$\left.\frac{\partial \pi}{\partial q_L}\right|_{(s^*,s^*)} = 0 \Rightarrow MRS\left(q^*, \omega + T - p^* - \frac{\tau}{2}\right) = MRT\left(q^*\right) \tag{2.8}$$

The first-order-condition from price means that the markup equals the acquisition cost parameter. The first order-condition from quality means that the indifferent consumer's marginal rate of substitution equals the marginal rate of transformation. The two equations imply that the equilibrium values are equal to the market values: $s^* = \left(p^\circ\left(\omega + T\right), q^\circ\left(\omega + T\right)\right)$. Since the market values for any given post-transfer income are unique, equilibrium must also be unique.

Voucher

Since equilibrium is non-zero quality is larger than zero and thus the first-order condition with respect to quality must hold, hence:

$$\left.\frac{\partial \pi}{\partial q_L}\right|_{(s^*,s^*)} = 0 \Rightarrow MRS\left(q^*, \omega + V - p^* - \frac{\tau}{2}\right)\frac{\mu\left(s^*\right)}{\tau} = MRT\left(q^*\right) \tag{2.9}$$

By the price floor proposition equilibrium price is at least as high as the voucher, i.e. $p^* \geq V$:

First, consider $p^* > V$. Then equation 2.5 applies and the first-order-condition with respect to price must hold:

$$\left.\frac{\partial \pi}{\partial p_L}\right|_{(s^*,s^*)} = 0 \Rightarrow \mu\left(s^*\right) = \tau \tag{2.10}$$

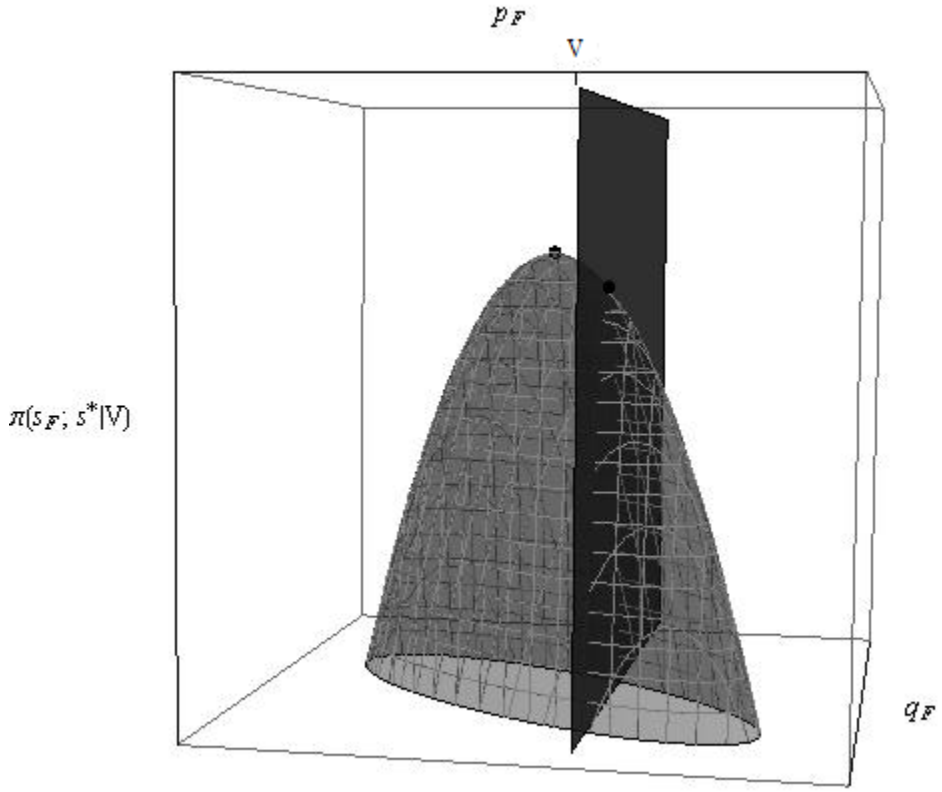


Figure 2.2: Proper voucher: Profit function given eq. behavior by $-F$.

But these two equations are just the definition of the market value for a post-transfer income of $I = \omega + V$, and equilibrium is given by the market values, $s^* = (p^\circ(\omega + V), q^\circ(\omega + V))$, and the voucher is improper.

Second, consider $p^* = V$. Then equilibrium is given by $p^* = V$ and equation 2.9. The appendix shows that equilibrium is unique. At this corner solution for price, equation 2.5 gives only the right-hand side derivative of profit with respect to price. The derivative must be negative or zero, since if it were positive, the firm could increase its profit by raising price. If the derivative equals zero equilibrium coincides with the market values and the voucher is improper. Else the derivative is negative and the voucher improper. The negative derivative implies that the markup is strictly

larger than the acquisition cost parameter τ:

$$\left.\frac{\partial \pi}{\partial^{+} p_L}\right|_{(s^*,s^*)} < 0 \Rightarrow \mu\left(s^*\right) > \tau \tag{2.11}$$

Figure 2.2 illustrates the case of an improper voucher.[6] A voucher is improper if and only if $V > p^{\circ}\left(\omega + V\right)$. The amount of the voucher V is indicated on the horizontal axis that gives p_F by the dark wall. Since the voucher can be understood as a price floor, firm F is effectively constrained to choose a price larger than V. The axis which seems to go into the page gives q_F, where the firm faces no constraint. The vertical axis gives $\pi(s_F; s^*|V)$, that is it gives the profit of firm F as a function of its choice variables p_F and q_F, if there is a voucher of V, and taking as given equilibrium behavior s^* by the other firm. The global maximum of the profit function is attained at a price that is smaller than the voucher or price floor and is thus not feasible. Therefore the firm chooses the constrained maximum as indicated in the figure. There the derivative of profit is zero in direction of the unconstrained variable, quality, but negative in direction of the bindingly constrained variable, price.

Quality floor

Consider the case where government sets a quality floor of $\underline{q} > 0$ in conjunction with a cash transfer of T. Since equilibrium is non-zero, the equilibrium price is larger than zero and thus in equilibrium the first order-condition with respect to price must hold:

$$\left.\frac{\partial \pi}{\partial p_L}\right|_{(s^*,s^*)} = 0 \Rightarrow \mu\left(s^*\right) = \tau \tag{2.12}$$

The profit-maximizing quality must be legal, i.e. either $q^* > \underline{q}$ or $q^* = \underline{q}$.

[6] The bottom left front corner of the box should be understood to be the origin for p_F and q_F but profit at that point is already positive. The figure should not be understood to ascertain that π is necessarily globally concave.

If $q^* > \underline{q}$, then the first-order-condition with respect to quality must hold:

$$\left.\frac{\partial \pi}{\partial q_L}\right|_{(s^*,s^*)} = 0 \Rightarrow MRS\left(q^*, \omega + T - p^* - \frac{\tau}{2}\right) = MRT\left(q^*\right) \tag{2.13}$$

But then the two equations are again just the definition of the market values, thus the equilibrium is given by the market values, that is $s^* = \left(p^\circ\left(\omega + T\right), q^\circ\left(\omega + T\right)\right)$. Therefore the quality floor is improper and equilibrium is unique.

Otherwise, we have $q^* = \underline{q}$, and thus equilibrium is given by $s^* = \left(c\left(\underline{q}\right) + \tau, \underline{q}\right)$. To see whether this is proper or improper, note that equation 2.6 gives the right-hand side derivative of profit with respect to quality. If it is equal to zero, equilibrium must equal the market values and thus the quality floor is improper. The derivative cannot be positive, as then the firm would do better by raising its quality. If the derivative is negative then the MRS is smaller than the MRT, so consumption of quality is distorted upward, and thus there is overconsumption of quality:

$$\left.\frac{\partial \pi}{\partial^+ q_L}\right|_{(s^*,s^*)} < 0 \Rightarrow MRS\left(q^*, \omega + T - p^* - \frac{\tau}{2}\right) < MRT\left(q^*\right) \tag{2.14}$$

2.5 Conclusion

Ever since Southworth (1945) it has been well-known that vouchers distort consumption. This chapter investigates the effect of an overlooked consequence of that consumption distortion: the fact that a distortionary or *proper* voucher reduces consumers' willingness to incur acquisition costs. The competition theorem (p.30) states that under variable acquisition costs proper vouchers lead to a decrease in competition compared to the market. Thus the competition attenuation occurs if and only if both of two conditions are met: the voucher must be proper and there must be variable acquisition costs.

Improper vouchers violate the first necessary condition and thus do not attenuate competition. An improper voucher is one that implements the same equilibrium as an unrestricted cash transfer of the same size does. From an applied perspective improper vouchers are of little interest anyway as presumably the whole reason for employing vouchers is to change consumption above and beyond some mere income effect.

Variable acquisition costs mean that there is some friction in the market which requires consumers to expend financial or non-financial resources in order to acquire the good. In the model a simple friction, travel cost, is used. Other examples would be bargaining, search, cognition or learning. In real world applications it is likely that at least one, if not indeed many of those, would be present. The acquisition costs must be such that in principle by varying them the consumer can obtain a more desirable combination of price and quality. This does not mean that acquisition expenditure will necessarily vary across equilibria in exercises of comparative statics. Indeed in this model in equilibrium consumers always purchase from the closest firm, and thus acquisition expenditure does not vary across the equilibria that are implemented by different policies, even though it is in principle variable. As often in economics it is the threat or off-equilibrium behavior that matters. Here in any equilibrium consumers purchase at the firm that is located closest to them, but the fact that they could purchase from the firm that is further away nevertheless matters. This point is worth bearing in mind when applying the model.

If there are no variable acquisition costs, the second necessary condition for competition attenuation is not met. Then a proper voucher still distorts consumption, but does not attenuate competition. The consumption distortion lowers the willingness to incur acquisition costs, but this is now inconsequential as there are no acquisition costs to be incurred or to be varied in the first place. In the model no aquisition costs correspond to an acquisition cost parameter of zero. The model has

the nice feature that firms' rents are continuous in the acquisition cost parameter. Firms' equilibrium rents turn out to be equal to the acquisition cost parameter and are zero in the limit as the acquisition cost parameter becomes zero. In that limit the model becomes a two-dimensional Bertrand model, and thus exhibits perfect competition. There even a proper voucher has no impact on competition: Observe that while it continues to de facto eliminate downward price competition, quality competition has now become perfect: an arbitrarily small increase in quality lets a firm capture the entire market, i.e. the *quality* elasticity of demand is infinite. Only in that case do vouchers entirely fulfill the hope that has been placed in them: they lead to an increase in spending on the good, competition ensures that quality will be higher, and there is no detrimental effect on competition.

From a welfare perspective the competition attenuation is distinct from the well-known "deadweight loss" of a voucher. The latter is just a utility loss for the consumer, as she would prefer an unrestricted cash transfer over the voucher. But if the voucher serves some policy objective, such as for example to address a positive consumption externality, the utility loss may not be a deadweight loss at all. Unlike the conventional utility loss, the competition attenuation result is a concern from the welfare perspective even if the voucher was designed to address such a positive externality.

The model compares vouchers to quality floors, and the latter turn out to be a benchmark: Quality floors can implement higher quality without attenuating competition in the slightest even when there are variable acquisition costs. This result is reminiscent of results in trade theory and environmental economics that state that if feasible direct targeting is better than indirect targeting. Therefore, if quality floors are feasible policymakers should employ quality floors, possibly in conjunction with cash transfers, rather than vouchers.

Interestingly, for education Mill (1859) suggested just such a policy. Mill proposed that government mandate that parents ensure that their children achieve a minimum educational level by each age and give poor parents unrestricted cash transfers to compensate them for the cost of that mandate. Unfortunately, the empirical reality teaches us that mechanisms which condition incentives on output lead to unintended consequences.

Quality floors can come as input or output quality floors. Output quality floors, which is what Mill called for, and which serve as the hypothetical benchmark in my model, seem to be empirically rare. That rarity finds a theoretical explanation in the mechanism design literature. Fixed and contractible output floors are necessarily narrow, and therefore invite dysfunctional responses as Holmstrom and Milgrom (1991) show. Perhaps the most prominent example of an output standard in education is the No Child Left Behind Act . In education output quality standards are often known as school accountability which means that school performance is measured by administering standardized tests. For a review of this literature see Figlio and Loeb (2011) which points to many of the dysfunctional responses to narrow standards as predicted by the mechanism design literature. Dysfunctional responses can range from teaching to the test (for evidence on this see Jacob, 2005) to outright cheating (see Jacob and Levitt, 2003). *Input* quality floors on the other hand, are common and occur in such varied forms as licenses, credentials, staff qualification or staffing level (e.g. classroom size) requirements. Input standards might do more to promote the interests of employees and incumbent firms, than to increase quality. For example, Hanushek et al. find that teacher credentials have no impact on learning outcomes (2005). Indeed maybe the potential of input standards to protect rents of incumbent firms and employees helps explain their commonality.

While quality floors pre-commit government to a well-defined and narrow purchasing or reward rule, vouchers allow government to decentralize the judgment

about quality by putting it into the hands of consumers (or their parents) themselves. Government does not have to commit to a particular measure or standard or procedures to determine whether they have been met and if not what the exact consequences are. In this understanding gaming a quality floor, as for example "teaching to the test" does, is much easier than gaming consumers. Presumably consumers also have better information about quality. If however government has certain pieces of information consumers do not have, then government can just make it available to consumers for free. The problem the model highlights is not one of exogenously given insufficient information by consumers. Rather consumers are endogenously at the margin not as much interested in higher quality as the government would want them to be. The model shows the precise sense in which this is true and caused by vouchers: an increase in quality that costs society a dollar to produce is valued at less than a dollar by consumers. This means that consumers are willing to spend less than a dollar of their own private resources on acquisition to acquire one dollar more of quality. While consumers behave optimally from their private perspective, they do not from a social one. The consumption distortion of a voucher, that is fully intended by government to increase consumption beyond the private optimum, is so-to-speak resisted by consumers who do not similarly distort their own private acquisition spending.

Policy conclusions that the model suggests depend on the particular application, what alternative policy instruments are feasible and if so at what cost. If no alternative policy instruments are available, then a rational policymaker upon learning about the detrimental effect of vouchers on competition would choose the amount of the voucher to be lower (and maybe cash transfers to be higher) than before. Essentially the policymaker has just learnt that there is a new trade-off: while a higher voucher amount leads to higher quality, it also lowers the level of competition more.

Another avenue to explore for policymakers is to pay particular attention to competition in markets where vouchers are predominant. Advertising, as a way that allows firms to shift acquisition costs to themselves, is of special interest in voucher markets, as consumer willingness to incur these is artificially reduced. In the same spirit firms should be allowed to pay and arrange a consumer's travel. Some private U.S. health insurers already do this and have started to offer travel and hotel accommodation to patients choosing to undergo surgery in places as far away from the U.S. as India.

A way to increase quality without vouchers is to make non-rival production inputs, such as knowledge, patents and licenses freely available. Policymakers may want to shift from public provision of private goods to public funding of non-rival inputs to these private goods. In education for example, private foundations or government could make textbooks available for the almost zero marginal cost of reproduction, through voluntary arrangements with the copyright holders such as prizes, competitions and buy-outs.

Future research should occur on both the theoretical and empirical fronts or combine both. Other sources of acquisition costs such as search, learning, or bargaining could be investigated to check the model for robustness; the intuition and mechanics of the model suggest that the competition attenuation will hold regardless of the precise source of the acquisition costs. Beyond that the model could be extended to dynamic settings to more formally discuss the implications for entry, industry structure and innovation. Empirical research is needed to quantify the effect of vouchers on competition, innovation and thus ultimately the long-run sign and size of the impact on welfare.

Appendix

Market values

***Claim* 2.1** (q° **well-defined**)

$f(q^\circ, I) = 0$, *where* f *is defined as*

$$f(q^\circ, I) \equiv MRS\left(q^\circ, I - \frac{3}{2}\tau - c(q^\circ)\right) - MRT(q^\circ)$$

defines a function q° *of* I *for all* $I > \frac{3}{2}\tau$.

Proof. Note that for $I > \frac{3}{2}\tau$ and $q^\circ > 0$ the MRS and MRT are real numbers, therefore $f(q^\circ, I)$ is a real number. Given an $I > \frac{3}{2}\tau$ observe that

(i) f is continuous in q° for $q^\circ > 0$,

(ii) for $q^\circ \to 0 : f(q^\circ) \to \infty$ (since $MRT(0)$ is finite by weak convexity of c and the MRS goes towards infinity as the indifference curves asymptote the axes).

(iii) for $q^\circ \to c^{-1}\left(I - \frac{3}{2}\tau\right)$: $MRS \to 0$, $MRT \to MRT\left(c^{-1}\left(I - \frac{3}{2}\tau\right)\right)$ and therefore there exists $q^\circ < c^{-1}\left(I - \frac{3}{2}\tau\right)$ such that $f(q^\circ, I) < 0$.

(i) - (iii) imply by the intermediate value theorem that there exists a q° with $0 < q^\circ < c^{-1}\left(I - \frac{3}{2}\tau\right)$ such that $f(q^\circ) = 0$. □

Note: Since f is a strictly decreasing function of q°, that q° is unique (MRS is strictly decreasing in q° since good 1 is not superior).

Symmetric and non-zero

***Claim* 2.2** (**positive demand**)

$d(s_L^*, s_R^*) > 0, d(s_R^*, s_L^*) > 0$*:*

Proof. First, suppose $d(s_L^*, s_R^*) = d(s_R^*, s_L^*) = 0$. Then firm L can deviate with $p_L = \omega + T + V - \tau$ and $q_L = \underline{q}$. Then markup $\mu(s_L) = \omega + T + V - \tau - c\left(\underline{q}\right) > 0$ by the assumption $\omega + T - \frac{3}{2}\tau > c\left(\underline{q}\right)$. Demand $d(s_L, s_R^*) = 1$ thus profit is strictly positive and the deviation profitable. Therefore it is not true that $d(s_L^*, s_R^*) = d(s_R^*, s_L^*) = 0$.

It remains to be shown that it is not true that one firm has zero demand and the other strictly positive demand. Suppose the contrary. Without loss of generality let the firm with positive demand be firm R, thus $d(s_R^*, s_L^*) > 0$ and $d(s_L^*, s_R^*) = 0$. In equilibrium a firm must be making non-negative profit, thus for R the markup must be non-negative, i.e. $\mu(s_R^*) \geq 0$.

(i) $\tau > 0$: Consider the deviation for firm L given by $s_L = \left(p_R^* + \frac{\tau}{2} d(s_R^*, s_L^*), q_R^*\right)$. Then demand is $d(s_L, s_R^*) \geq \frac{1}{4}$, and thus profit is $\pi(s_L, s_R^*) \geq \frac{1}{4}\left(\mu(s_R^*) + \frac{\tau}{2} d(s_R^*, s_L^*)\right) > 0$, contradicting the supposition.

(ii) $\tau = 0$, and thus $d(s_R^*, s_L^*) = 1$ and $u(q_R^*, \omega + T + V - p_R^*) > u(q_L^*, \omega + T + V - p_L^*)$. But then firm R could deviate with $(p_R^* + \varepsilon, q_R^*)$ for some small ε, keep its demand unchanged and thus make higher profits. □

***Claim* 2.3** (**positive consumption of good 1**)

$d(s_L^*, s_R^*) + d(s_R^*, s_L^*) = 1$

Proof. For $\tau = 0$ this is immediately implied by positive demand.

$\tau > 0$: Suppose to the contrary that $d(s_L^*, s_R^*) + d(s_R^*, s_L^*) < 1$. Without loss of generality assume in this step that $d(s_L^*, s_R^*) < d(s_R^*, s_L^*)$, and therefore $d(s_L^*, s_R^*) < \frac{1}{2}$. As $i = d(s_L^*, s_R^*)$ is the last consumer able to afford buying from L, firm L's price must be such that her consumption of the numeraire good equals zero: $\omega + T + \min\{V, p_L^*\} - p_L^* - d^*(s_L^*, s_R^*)\tau = 0$. Since $\omega > \tau$ and $T \geqslant 0$ this implies that $p_L^* > V$. Therefore the price is given by $p_L^* = \omega + T + V - d(s_L^*, s_R^*)\tau$ and thus $\pi(s_L^*, s_R^*) = d(s_L^*, s_R^*) \cdot (\omega + T + V - d(s_L^*, s_R^*)\tau - c(q_L^*))$. Consider the deviation: $p_L = \omega + T + V - \tau(1 - d_L^*)$ and $q_L = \underline{q}$. Then $\pi(s_L, s_R^*) = (1 - d_L^*)(\omega + T + V - \tau(1 - d_L^*) - c(q_L))$. Thus the deviation is profitable as $(1 - d_L^*)(\omega + T + V - \tau(1 - d_L^*) - c(q_L)) > d(s_L^*, s_R^*)(\omega + T + V - \tau d_L^* - c(q_L^*))$. (In the case of $\underline{q} > 0$ this is true since by assumption a fortiori $\omega - \tau > c(\underline{q})$). □

***Lemma* 2.2** (**symmetric**)

If (s_L^*, s_R^*) is an equilibrium then it is symmetric: $s_L^* = s_R^*$.

Proof. See claims below. □

***Claim* 2.4** (quality symmetric)

$q_L^* = q_R^*$.

Proof. Suppose to the contrary $q_L^* \neq q_R^*$. Since all consumers buy good 1, and each firm has non-zero demand, there exists a consumer who is indifferent between purchasing from either firm, and this indifferent or swing consumer is $i = d(s_L^*, s_R^*) = 1 - d(s_R^*, s_L^*)$. This swing consumer is indifferent between the bundle that results from her purchasing from firm L, $\left(q_L^*, \omega + T + V - \tau\widetilde{i} - p_L^*\right)$ and the one that results from her purchasing from firm R, $\left(q_R^*, \omega + T + V - \tau\left(1 - \widetilde{i}\right) - p_R^*\right)$. By the supposition these bundles are not identical, thus strong convexity implies that the swing consumer strongly prefers their average $\left(\frac{1}{2}q_L^* + \frac{1}{2}q_R^*, \omega + T + V - \frac{\tau}{2} - \left(\frac{1}{2}p_L + \frac{1}{2}p_R^*\right)\right)$ to the bundles themselves. Thus if a firm deviates with s_F which is defined by price and quality such that the swing consumer ends up with the average bundle if purchasing from it, then the deviating firm's demand must be strictly bigger than in the equilibrium. Therefore its markup under the deviation must be strictly smaller than its equilibrium markup, that is $\mu(s_F) < \mu(s_F^*)$. For firm L this deviation is given by $p_L = \frac{1}{2}p_L^* + \frac{1}{2}p_R^* + \tau\left(\frac{1}{2} - d(s_L^*, s_R^*)\right)$ and $q_L = \frac{1}{2}q_L^* + \frac{1}{2}q_R^*$, while for firm R deviation is $p_R = \frac{1}{2}p_L^* + \frac{1}{2}p_R^* + \tau\left(d(s_L^*, s_R^*) - \frac{1}{2}\right)$ and $q_R = \frac{1}{2}q_L^* + \frac{1}{2}q_R^*$.

$$\mu(s_L) < \mu(s_L^*) \Rightarrow \frac{p_R^* - p_L^*}{2} + c(q_L^*) - c\left(\frac{q_L^* + q_R^*}{2}\right) < \tau\left(d(s_L^*, s_R^*) - \frac{1}{2}\right)$$
$$\mu(s_R^*) < \mu(s_R) \Rightarrow \tau\left(d(s_L^*, s_R^*) - \frac{1}{2}\right) < \frac{p_R^* - p_L^*}{2} - c(q_R^*) + c\left(\frac{q_L^* + q_R^*}{2}\right)$$
$$\Rightarrow c\left(\frac{q_L^* + q_R^*}{2}\right) > \frac{1}{2}c(q_L^*) + \frac{1}{2}c(q_R^*)$$

But this contradicts convexity of c, and thus the supposition is wrong. □

***Claim* 2.5** (price symmetric)

$p_L^* = p_R^*$.

Proof. Suppose not, and without loss of generality suppose that $p_L^* < p_R^*$. Since all consumers buy good 1, and each firm has non-zero demand consumer $i = d(s_L^*, s_R^*) = 1 - d(s_R^*, s_L^*)$ is indifferent between purchasing from either firm. By claim 2.4 $q_L^* = q_R^* = q^*$,

therefore for the indifferent consumer the numeraire consumption resulting from purchasing at either firm must be the same, thus:

$$\omega + T + V - \tau d\left(s_L^*, s_R^*\right) - p_L^* = \omega + T + V - \tau\left(1 - d\left(s_L^*, s_R^*\right)\right) - p_R^*$$
$$\Rightarrow p_R^* - p_L^* = \tau\left(2d\left(s_L^*, s_R^*\right) - 1\right)$$
$$\Rightarrow d\left(s_L^*, s_R^*\right) > \frac{1}{2} \text{ and } \mu\left(s_R^*\right) = \mu\left(s_L^*\right) + \tau\left(2d\left(s_L^*, s_R^*\right) - 1\right)$$

Firm R could deviate by playing s_L^*, and this deviation cannot be profitable, hence:

$$\pi\left(s_L^*, s_L^*\right) \leq \pi\left(s_R^*, s_L^*\right)$$
$$d\left(s_L^*, s_L^*\right)\mu\left(s_L^*\right) \leq \left(1 - d\left(s_L^*, s_R^*\right)\right)\mu\left(s_R^*\right)$$

Since consumer $i = \frac{1}{2}$ can afford price p_L^*, as $d\left(s_L^*, s_R^*\right) > \frac{1}{2}$ we know that $d\left(s_L^*, s_L^*\right) = \frac{1}{2}$[7], substituting this and the markup yields:

$$\frac{1}{2}\mu\left(s_L^*\right) \leq \left(1 - d\left(s_L^*, s_R^*\right)\right)\left(\mu\left(s_L^*\right) + \tau\left(2d\left(s_L^*, s_R^*\right) - 1\right)\right)$$
$$\mu\left(s_L^*\right) \leq 2\tau\left(1 - d\left(s_L^*, s_R^*\right)\right) \Rightarrow \mu\left(s_L^*\right) < \tau$$

The latter implies $\mu\left(s_L^*\right) < \tau$, and since $\mu\left(s_R^*\right) = \mu\left(s_L^*\right) + \tau\left(2d\left(s_L^*, s_R^*\right) - 1\right)$, moreover $\mu\left(s_R^*\right) \leq \tau$. Now consider two cases, either the numeraire consumption of the indifferent consumer in equilibrium is positive or it is zero:

Case (i): Numeraire consumption is positive:

If a firm raises price it loses 2τ consumers, i.e. $\frac{\partial d\left(s_L^*, s_R^*\right)}{\partial p_L} = -\frac{1}{2\tau}$ and thus:

$$\frac{\partial \pi\left(s_L^*, s_R^*\right)}{\partial p_L} = \frac{\partial d\left(s_L^*, s_R^*\right)}{\partial p_L}\mu\left(s_L^*\right) + d\left(s_L^*, s_R^*\right) = -\frac{1}{2\tau}\mu\left(s_L^*\right) + d\left(s_L^*, s_R^*\right) = 0$$
$$\Rightarrow \frac{1}{2\tau}\mu\left(s_L^*\right) = d\left(s_L^*, s_R^*\right) > \frac{1}{2} \Rightarrow \mu\left(s_L^*\right) > \tau$$

[7]Note that this argument would not work for $d\left(s_R^*, s_R^*\right) = \frac{1}{2}$.

Which contradicts $\mu(s_L^*) < \tau$.

Case (ii): Numeraire consumption is zero:

This means

$$\omega + T + V - \tau d(s_L^*, s_R^*) - p_L^* = 0$$

$$\Rightarrow p_L^* = \omega + T + V - \tau d(s_L^*, s_R^*)$$

$$\Rightarrow \ p_R^* = \omega + T + V - \tau(1 - d(s_L^*, s_R^*))$$

Then a firm does not lose consumers by lowering quality, but doing so still increases the markup, thus quality must already be at the minimum: $q^* = \underline{q}$.

$$\mu(s_R^*) = \omega + T + V - \tau(1 - d(s_L^*, s_R^*)) - c(\underline{q}) > \omega + T + V - \frac{1}{2}\tau - c(\underline{q}) > \tau$$

□

Which contradicts the earlier observation that $\mu(s_R^*) \leq \tau$.

***Claim* 2.6** (**positive numeraire**)

$\omega + T + V - p^* - \frac{1}{2}\tau > 0$

Proof. By claim 2.3 the equilibrium price satisfies $\omega + T + V - p^* - \frac{1}{2}\tau \geq 0$, thus it only remains to be shown that shown that $\omega + T + V - p^* - \frac{1}{2}\tau \neq 0$. To the contrary suppose $\omega + T + V - p^* - \frac{1}{2}\tau = 0$. Then $p^* = \omega + T + V - \frac{1}{2}\tau$, therefore $\pi(s^*; s^*) = \frac{1}{2}(p^* - c(q^*))$. Consider the deviation $p_L = p^* - \frac{1}{2}\tau$ and $q_L = \underline{q}$ which means that demand will be 1 and

profit $\pi(s_L; s^*) = p^* - \frac{1}{2}\tau - c\left(\underline{q}\right)$.

$$\pi(s_L, s^*) \leq \pi(s^*, s^*)$$
$$p^* - \frac{1}{2}\tau - c\left(\underline{q}\right) \leq \frac{1}{2}\left(p^* - c\left(q^*\right)\right)$$
$$\omega + T + V - \frac{3}{2}\tau \leq c\left(\underline{q}\right) - \left(c\left(q^*\right) - c\left(\underline{q}\right)\right)$$
$$\Rightarrow \omega + T + V - \frac{3}{2}\tau \leq c\left(\underline{q}\right)$$

But this contradicts $\omega + T - \frac{3}{2}\tau > c\left(\underline{q}\right)$. □

***Claim* 2.7** (**positive quality**)

$q^* > 0$

Proof. If $\underline{q} > 0$ then trivially $q^* > 0$.

Else $\underline{q} = 0$, and suppose to the contrary of the claim that $q^* = 0$. By the prior claim $p^* < \omega + T + V - \frac{1}{2}\tau$, and therefore $\pi(s^*, s^*) < \frac{1}{2}\left(\omega + T + V - \frac{1}{2}\tau\right)$. Consider the deviation $p_L = \omega + T + V - \tau$ and $q_L = c^{-1}\left(\frac{1}{2}\left(\omega + T + V - \frac{3}{2}\tau\right)\right)$. Then demand is 1, and thus profit

$$\pi(s_L, s^*) = \frac{1}{2}\left(\omega + T + V - \frac{1}{2}\tau\right) > \pi(s^*; s^*)$$

□

***Lemma* 2.3** (**non-zero**)

If (s^*, s^*) is an equilibrium then it is non-zero.

Proof. Follows from the previous claims. □

Equivalence

***Claim* 2.8** (**demand policy invariant**)

For all (s_L, s_R) with $p_L, p_R \geq \underline{p}$ and $q_L, q_R \geq \underline{q}$: If $V = \underline{p} = T$ then $d(s_L, s_R|V) = d\left(s_L, s_R|\underline{p},T\right)$.

Proof. If a consumer buys from firm F given a voucher of V, she consumes quality q_F and has a numeraire consumption of $\omega + \min\{V, p_F\} - \tau i_F - p_F = \omega + V - \tau i_F - p_F$. If a consumer buys from firm F given a price floor of $\underline{p} = V$ and cash transfer T, she consumes quality q_F and has a numeraire consumption of $\omega + T - \tau i_F - p_F = \omega + V - \tau i_F - p_F$. Therefore the bundles available for purchase, i.e. consumers' choice sets are identical in both situations, and thus their unique choices must coincide and thus demand must be the same. □

***Claim* 2.9** (**perfectly price inelastic**)

For any two strategies s_F and s'_F that satisfy p_F, $p'_F \leq V$ and $q_F = q'_F$ and all strategies s_{-F}: $d(s_F, s_{-F}|V) = d\left(s'_F, s_{-F}|V\right)$.

Proof. For all consumers i numeraire consumption when buying at a price lower or equal to V is $\omega - \tau i_F$. Therefore a consumer's utility when buying from F is given by $u(q_F, \omega - \tau i_F)$ regardless of whether firm F has chosen s_F or s'_F and therefore $d(s_F, s_{-F}) = d\left(s'_F, s_{-F}\right)$. □

***Claim* 2.10** (**de-facto price floor**)

If $d\left(s^*_F, s^*_{-F}|V\right) > 0$ then $p^*_F \geq V$.

Proof. Suppose to the contrary that $p^*_F < V$: Consider the deviation $s_F = (V, q^*_F)$ which by claim 2.9 has the same demand as s^*_F: $d\left(s_F, s^*_{-F}|V\right) = d\left(s^*_F, s^*_{-F}|V\right)$. Since $\mu(s_F) > \mu(s^*_F)$ it is the case that $\pi\left(s_F, s^*_{-F}|V\right) > \pi\left(s^*_F, s^*_{-F}|V\right)$, which contradicts the claim that $\left(s^*_F, s^*_{-F}\right)$ is an equilibrium. □

***Proposition* 2.2** (**equivalence**)

See page 29.

Proof.

I. If (s^*_L, s^*_R) is implemented by $(\underline{p}, T)$ then V implements (s^*_L, s^*_R):

Suppose not, then there exists a firm, without loss of generality let it be firm L, and a strategy s_L such that $\pi\left(s_L, s_R^*|\underline{p},T\right) \leq \pi\left(s_L^*, s_R^*|\underline{p},T\right)$, but $\pi\left(s_L, s_R^*|V\right) > \pi\left(s_L^*, s_R^*|V\right)$.

$$\pi\left(s_L, s_R^*|\underline{p},T\right) \leq \pi\left(s_L^*, s_R^*|\underline{p},T\right)$$
$$d\left(s_L, s_R^*|\underline{p},T\right)\mu\left(s_L\right) \leq d\left(s_L^*, s_R^*|\underline{p},T\right)\mu(s_L^*)$$

By claim 2.8 (demand invariant) therefore:

$$d\left(s_L, s_R^*|V\right)\mu\left(s_L\right) \leq d\left(s_L^*, s_R^*|V\right)\mu(s_L^*)$$

But the latter is contradicted by:

$$\pi\left(s_L, s_R^*|V\right) > \pi\left(s_L^*, s_R^*|V\right)$$
$$d\left(s_L, s_R^*|V\right)\mu\left(s_L\right) > d\left(s_L^*, s_R^*|V\right)\mu\left(s_L^*\right)$$

Therefore V implements (s_L^*, s_R^*).

II. If (s_L^*, s_R^*) is implemented by V, then $(\underline{p}, T)$ implements (s_L^*, s_R^*):

By claim 2.10 (de facto price floor) $p_L^*, p_R^* \geq V$, thus p_L^*, p_R^* are legal under the price floor $\underline{p} = V$. Thus if we suppose that $(\underline{p}, T)$ does not implement (s_L^*, s_R^*), then there exists a firm, without loss of generality let it be firm L, and a strategy s_L such that $\pi\left(s_L, s_R^*|\underline{p},T\right) > \pi\left(s_L^*, s_R^*|\underline{p},T\right)$ and $\pi\left(s_L, s_R^*|V\right) \leq \pi\left(s_L^*, s_R^*|V\right)$. By lemma 2.3 (non-zero) $d\left(s_L^*, s_R^*\right) > 0$, and therefore by claim 2.9 (price inelastic below V) $p_F^* \geq V$, and therefore $p_F^* \geq \underline{p}$.

$$\pi\left(s_L, s_R^*|\underline{p},T\right) > \pi\left(s_L^*, s_R^*|\underline{p},T\right)$$
$$d\left(s_L, s_R^*|\underline{p},T\right)\mu\left(s_L\right) > d\left(s_L^*, s_R^*|\underline{p},T\right)\mu\left(s_L^*\right)$$

By claim 2.8 (demand invariant) therefore:

$$d\left(s_L, s_R^*|V\right)\mu\left(s_L\right) > d\left(s_L^*, s_R^*|V\right)\mu\left(s_L^*\right)$$

But this contradicts the following:

$$\pi\left(s_L, s_R^*|V\right) \leq \pi\left(s_L^*, s_R^*|V\right)$$
$$d\left(s_L, s_R^*|V\right)\mu\left(s_L\right) \leq d\left(s_L^*, s_R^*|V\right)\mu(s_L^*)$$

□

Perfect Competition ($\tau = 0$)

Proof (Lemma zero markup). Suppose to the contrary $\mu\left(s^*\right) > 0$. Then equilibrium profit is, by symmetry and non-zero, $\pi\left(s^*, s^*\right) = \frac{1}{2}\mu\left(s^*\right) > 0$. Consider the deviation $s = \left(p^*, c^{-1}\left(c\left(q^*\right) + \frac{\mu(s^*)}{3}\right)\right)$. Then $d\left(s, s^*\right) = 1$, thus profit is $\pi\left(s, s^*\right) = \frac{2}{3}\mu\left(s^*\right) > \pi\left(s^*, s^*\right)$, contradicting the supposition. □

Imperfect Competition ($\tau > 0$)

***Claim* 2.11** (**proper voucher implements uniquely**)
For $\tau > 0$: Let the function k of q be given by $k\left(q\right) = MRS\left(q, \omega + T - V - \frac{\tau}{2}\right)\frac{V - c(q)}{\tau} - c_q\left(q\right)$. For any proper V there exists a unique q such that $k\left(q\right) = 0$.

Proof. Observe that (i) k is continuous, (ii) $q \to 0 : k(q) \to \infty$, (iii) for all $q \geqslant c^{-1}\left(V\right) : k(c^{-1}\left(V\right)) < 0$. Thus by the intermediate value theorem there exists at least one q such that $k\left(q\right) = 0$. Note that k is a strictly decreasing function of q, establishes that there is exactly one q with $k\left(q\right) = 0$. □

Existence

It is possible to construct counterexamples to existence under these very general assumptions. One such counterexample can be found in the appendix 2.1. However holding the

utility and cost functions fixed, as well as income ω, it is likely true that for sufficiently small τ equilibrium exists as the following conjecture claims:

***Conjecture* 2.1** (**existence**)

Given a utility function u, a cost function c, income ω, and a policy there exists $\overline{\tau} > 0$ such that for all τ with $0 \leq \tau < \overline{\tau}$ an equilibrium exists.

If true the conjecture implies that for $\tau = 0$ equilibrium always exists, a claim which has been shown to be true.

The assumptions made for the competition theorem are not sufficient to ensure existence. The Nash-theorem is not applicable as firms' profit functions do not have to be quasi-concave[8]. Quasi-concavity of the profit function is problematic due to the two-dimensional strategy space. The candidate equilibrium as defined by the necessary conditions is not necessarily a global optimum. The following gives a counterexample to existence.

***Example* 2.1** (**non-existence**)

$u(x_1, x_2) = x_1 x_2$, $c(q_F) = q_F$, $\tau = \frac{2}{3}$ $\omega = \frac{25}{24}$, *and market policy with* $T = 0$.

The necessary conditions are $MRS\left(q^*, \omega - p^{**} - \frac{\tau}{2}\right) = c_q(q^*)$ and $p^* - c(q^*) = \tau$. Thus $q^* = \frac{1}{48}$ and $p^* = \frac{33}{48}$, and the numeraire consumption of the indifferent consumer is $x_2 = \frac{1}{48}$, markup is $\mu(s^*) = \frac{2}{3}$, and a firm's profit is $\pi(s^*, s^*) = \frac{1}{3}$. Consider the deviation $s_L = \left(\frac{13}{20}, \frac{3}{100}\right)$. Then markup is $\mu(s_L) = \frac{62}{100}$, thus the deviation is profitable if and only if if the firm's demand is larger than $\frac{\mu\left(s_L^*\right)}{2\mu(s_L)} = \frac{50}{93}$. Thus it suffices to show that consumer $i = \frac{13}{24} > \frac{50}{93}$ strictly prefers to purchase from firm L rather than R. Purchasing from L consumer$\frac{13}{24}$ has utility $u\left(\frac{13}{20}, \frac{11}{30 \cdot 12}\right) = \frac{13 \cdot 11}{50 \cdot 12^2}$, which is larger than the utility of purchasing from R, $u\left(\frac{1}{48}, \frac{7}{144}\right) = \frac{7}{4 \cdot 12^3}$.

[8]Continuity is satisfied. Compactness is not satisfied, but is not the reason for the possible non-existence.

References

Baumol, W. J. and W. G. Bowen (1965). On the performing arts: the anatomy of their economic problems. *American Economic Review 55*(1/2), 495–502.

BEA (2011). General government expenditure by function, Table 1100 for 2010. Report, Bureau of Economic Analysis, Washington, D.C.

Blomquist, S. and V. Christiansen (1999). The political economy of publicly provided private goods. *Journal of Public Economics 73*(1), 31–54.

Bradford, D. F. and D. N. Shaviro (2000). The economics of vouchers. In C. E. Steuerle (Ed.), *Vouchers and the provision of public services*, Chapter 2, pp. 40–91. Washington, DC: Brookings Institution Press.

Buchanan, J. M. (1975). *The Samaritan's Dilemma*, pp. 71–85. Russell Sage Foundation.

Congressional Budget Office (2011). CBO's long-term budget outlook. Report, Congressional Budget Office, Washington, D.C.

Currie, J. and F. Gahvari (2008). Transfers in cash and in kind: theory meets the data. *Journal of Economic Literature 46*(2), 333–83.

Epple, D., E. Newlon, and R. Romano (2002). Ability tracking, school competition, and the distribution of educational benefits. *Journal of Public Economics 83*(1), 1–48.

Epple, D. and R. E. Romano (1996). Public provision of private goods. *Journal of Political Economy 104*(1), 57–84.

Figlio, D. and S. Loeb (2011). School accountability. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbooks in Economics: Economics of Education*, Chapter 8, pp. 383–421. North Holland: Elsevier.

Finkelstein, A. (2007). The aggregate effects of health insurance: evidence from the introduction of medicare. *Quarterly Journal of Economics 122*(1), 1–37.

Fogel, R. W. (2004). Health, nutrition, and economic growth. *Economic Development and Cultural Change 52*, 643–58.

Friedman, M. (1955). The role of government in education. In R. A. Solo (Ed.), *Economics and the Public Interest*. New Brunswick, NJ: Rutgers University Press.

Hanson, R. (2008). Showing that you care: the evolution of health altruism. *Medical Hypotheses 70*(4), 724–742.

Hanushek, E. A., J. F. Kain, D. M. O'Brien, and S. G. Rivkin (2005). The market for teacher quality. Working Paper 11154, NBER.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization 7*, 24–52.

Jacob, B. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the chicago public schools. *Journal of Public Economics 89*(5), 761–96.

Jacob, B. A. and S. D. Levitt (2003). Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics 118*(3), 843–77.

Mill, J. S. (1859). *On Liberty*. London: John W. Parker and Son.

Mulligan, C. B. and T. J. Philipson (2000). Merit motives and government intervention: public finance in reverse. Working Paper 7698, NBER.

Musgrave, R. A. (1959). *The Theory of Public Finance: A Study in Public Economy*. McGraw-Hill New York.

Nechyba, T. J. (1996). Public school finance in a general equilibrium Tiebout world: equalization programs, peer effects and private school vouchers. Working Paper 5642, NBER.

Nichols, A. L. and R. J. Zeckhauser (1982). Targeting transfers through restrictions on recipients. *American Economic Review 72*(2), 372–377.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan and T. Cadell.

Southworth, H. M. (1945). The economics of public measures to subsidize food consumption. *Journal of Farm Economics 27*(1), 38–66.

Summers, L. H. (1989). Some simple economics of mandated benefits. *American Economic Review 79*(2), 177–83.

Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy 64*(5), 416–24.

v. Hayek, F. A. (1968). Der Wettbewerb als Entdeckungsverfahren (Competition as a Discovery Procedure). Kieler Vorträge, Institut für Weltwirtschaft Kiel.

Whitmore, D. (2002). *What Are Food Stamps Worth?* Working Paper. Princeton, NJ: Industrial Relations Section, Department of Economics, Princeton University.

Chapter 3

Normal goods

Abstract: A good is normal if its demand increases in income. Thus normality of a good is conventionally defined as a property of the demand function. This chapter defines normality as a local property of a preference relation. The exercise is relatively straightforward as long as convexity of the preference relation continues to be assumed. Without convexity I define normality of good 1 as follows: the slope of the indifference curve or the MRS increases as good 2 is increased. Or, equivalently, the distance between two indifference curves, decreases as good 1 is increased. This new, generalized definition of normality is useful when the dimension of good 1 is not cardinal. Chapter 4, using this new definition, shows that normality of good 1, unlike convexity, is invariant to order-preserving transformations of dimension 1. Inferiority can analogously be defined: the distance between two indifference curves increases as good 1 is increased. If a good is on the border between being normal and inferior, I propose to call it a *constant* good. If good 1 is constant then the distance between two indifference curves is constant. This means that the indifference curves are parallel, which is a well-known characteristic of quasilinear preferences. Whenever a good is quasi-linear the other good is constant, one might say constant goods form the entourage of a quasilinear good. Finally I show that if a preference is normal in *all* goods then the preference is convex.

3.1 Introduction

Normality of a good is a common assumption in applications as diverse as health economics, insurance, behavioral economics, labor economics, political economy, welfare economics, and the public or private provision of public goods[1]. Economists call a good normal if its demand increases as income increases (e.g. Varian 1992, p. 117). If demand for a good decreases in income, it is called inferior. If demand is constant as income increases there, being no term yet in the literature to the best of my knowledge, I call the good *constant*. Conventionally, normality is viewed as a property of the demand function, which is not in line with the standard approach in choice theory, which is to define such properties on the preference relation. But viewed from that perspective, normality is a local property of a preference relation with respect to a particular good. Monotonicity is similar in this respect, as it too is a is a local property of a preference with respect to a particular good. For an introduction to and overview of that choice theoretic approach see Kreps (1988). In this chapter the analysis is confined to the case of two goods.

To ensure existence of a demand function convexity is usually assumed. A definition on the preference rather than the demand function allows me to drop the assumption of convexity, and to investigate if normality has a meaning independently of convexity.

The definition I propose means that if a preference is normal in both goods then it is convex. A preference that is inferior in one good and normal in the other may or may not be convex. This result, that normality in both goods is a stronger property than convexity, can be seen as a novel justification for assuming convexity.

[1]See for example Gertler et al., 1987 (health economics), Cummins and Mahul, 2004 (insurance), Nelson, 2001 (behavioral economics), Green and Kahn, 1983 (labor economics), Besley and Coate, 1991 (political economy), Small and Rosen, 1981 (welfare economics), Wilson, 1991 (public provision of public goods), or Bergstrom et al., 1986 (private provision of public goods).

The proposed definition is a generalization of the existing one: for convex preferences it coincides with the existing one. The generalization is useful when the dimension of the good is not cardinal. Cardinality of a dimension is a natural assumption if the dimension measures quantity, but it is not natural for many other attributes a dimension may measure, such as quality of a service, size of a discrete physical good or strength of a political preference. Chapter 4 shows that normality of a good is invariant to order-preserving transformations of its own dimension. That statement is only well-defined due to the generalized definition of normality this chapter develops, as convexity may be lost under order-preserving transformations of dimension 1.

3.2 Two characterizations of normality

In the most common usage in economics good 1 is called normal if for any positive prices p and income ω demand $x^*(p, \omega)$ increases in income.[2] In that view normality is a property of the demand function. Indirectly though it should then also be a property of the underlying preference relation as well. To make this a local property of a preference any bundle should be demanded for some budget set. More precisely for every interior bundle x^0 there needs to exist a budget set such that x^0 is demanded. To ensure this assume that the preference relation is strongly monotone and strongly convex. For the slope of the indifference curves or the MRS to exist, assume that the preference relation is differentiable. In other words the preference admits a differentiable utility representation. The formal definition of differentiability of a preference itself is due to Rubinstein (2010). Consider this geometrically at x^0: By differentiability there exists a unique tangent through x^0. Normality means that when we shift

[2] Kreps (1990, p.49) defines a normal good as one whose demand decreases in *price*, this is not the usage followed here.

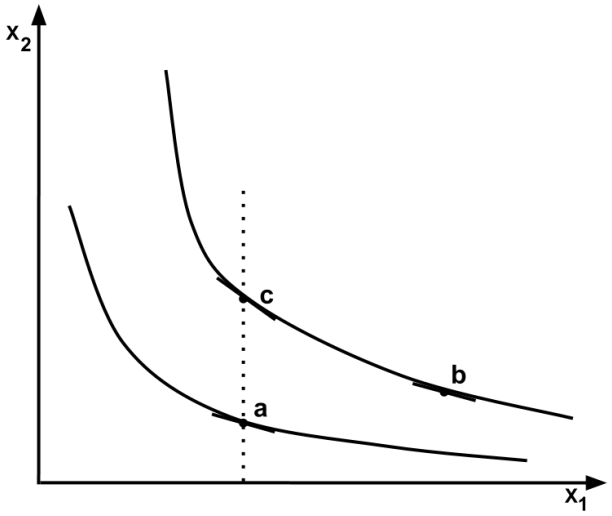
the tangent "outwards" (increase income), the point of tangency moves to the "right" (x_1^* increases). This is the tangency-characterization of normality.

From the tangency-characterization it is not obvious what normality means for the shape of the indifference curves. It turns out that there is an equivalent characterization of normality, the slope-characterization: the slope of the indifference curves gets steeper as x_2 is increased. Figure 3.1a illustrates why the indifference curves must get steeper: Given p, ω^0 the consumer chooses the bundle $a \equiv x^*(p, \omega^0)$. Holding prices fixed and increasing income to $\omega^1 > \omega^0$ is an outward parallel shift of the budget line, and normality of good 1 implies that the new choice $b \equiv x^*(p, \omega^1)$ satisfies $x_1^*(p, \omega^1) > x_1^*(p, \omega^0)$. By construction the slopes of the indifference curves at a and at b are the same. Moving from b along the indifference curve until the first coordinate is the same as that of the original point a, we reach point c. Convexity of the indifference curve means that the slope at c is steeper than at b, and therefore a.Note that getting from the tangency-characterization to the slope-characterization makes use of convexity of the preference, and indeed the equivalence of the tangency- and slope-conditions only holds under convexity.

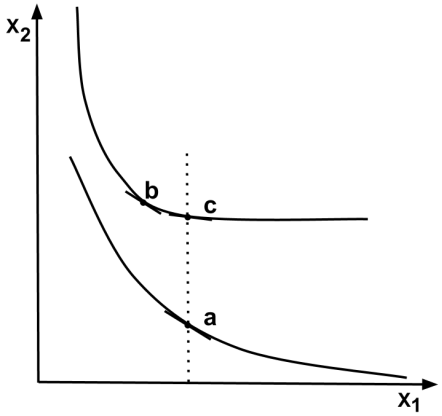
Figure 3.1b illustrates the analogous consideration for an inferior good: the point of tangency moves to the "left", from a to b, and the indifference curves get flatter as x_2 increases. Figure 3.1c illustrates the case of a good whose demand is constant in income: the tangency point does not move left or right, and the slope of the indifference curves stays constant as x_2 is increased. I call such a good a constant good. Formally define a constant good indirectly as a good that is both weakly normal and weakly inferior:

***Definition* 3.1**

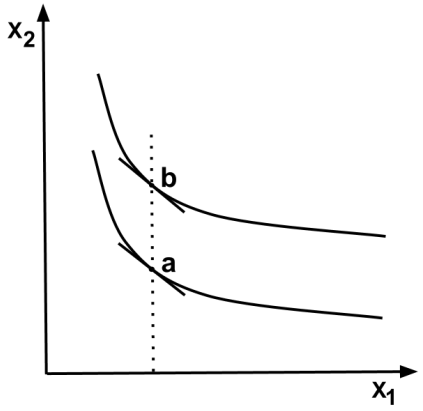
If a preference is normal and inferior in good 1 at x^0, then it is called constant in good 1 at x^0.



(a) Normal good



(b) Inferior good



(c) Constant good

Figure 3.1: Change of the slope of the indifference curve as x_2 increases.

If good 1 is a constant good, then the indifference curves are parallel-shifts of one another in direction of the x_2-axis. This means that good 2 is quasi-linear.

Now derive analytical versions of the characterizations. Assume that the preference is represented by a twice continuously differentiable utility function $u : X \subseteq \mathbb{R}^2_{++} \to \mathbb{R}$ that satisfies $u_1 > 0$, $u_2 > 0$ (essentially strong monotonicity) and $2u_{12}u_1u_2 - u_{11}u_2^2 - u_{22}u_1^2 > 0$ (essentially strong quasiconcavity). Then the derivative of demand for good 1 with respect to (normalized) income exists, and can be expressed as $\frac{\partial x_1^*}{\partial \frac{\omega}{p_1}} = \dfrac{u_{12}(x^*)\,u_2(x^*) - u_{22}(x^*)\,u_1(x^*)}{2u_{12}(x^*)\,u_1(x^*)\,u_2(x^*) - u_{11}(x^*)\,u_2^2(x^*) - u_{22}(x^*)\,u_1^2(x^*)} u_1(x^*)$.

Thus the tangency characterization says that a preference is normal in good 1 at x^0 if $\frac{u_{12}(x^0)u_2(x^0) - u_{22}(x^0)u_1(x^0)}{2u_{12}(x^0)u_1(x^0)u_2(x^0) - u_{11}(x^*)u_2^2(x^*) - u_{22}(x^*)u_1^2(x^*)} u_1(x^*) \geq 0$.

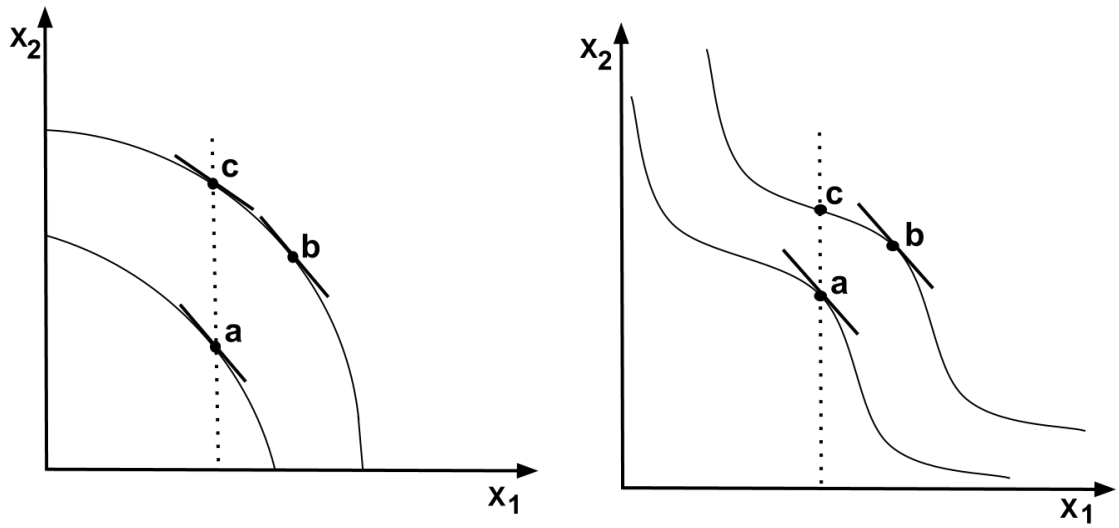
The slope characterization in its analytical form simply means that the preference is normal in good 1 at x^0, if the $MRS(x) \equiv \frac{u_1(x)}{u_2(x)}$ is increasing in x_2 at x^0, or since under our assumptions the derivative of the MRS exists, $\frac{\partial MRS}{\partial x_2}(x^0) = \frac{u_{12}(x^0)u_2(x^0) - u_{22}(x^0)u_1(x^0)}{(u_2(x^0))^2} \geq 0$.

These characterizations are not original to this paper. For example Leroux (1987, p.196) gives a condition for both goods being normal. Shitovitz and Spiegel (2001) define "ordinal normality" of a preference by the second characterization. Bilancini and Boncinelli (2010) prove that a good is normal if and only if this condition holds. They give analogous conditions for the case of more than 2 goods. All of these authors assume strongly convex preferences.

To see that the analytical versions of the tangency- and slope-characterizations are equivalent note that the denominator in the tangency characterization is the determinant of the bordered Hessian of u at x^0 and therefore strictly positive by strict quasi-concavity of u.

3.3 Defining normality without convexity

The conventional definition of normality employs and thus requires existence of the demand function. Consequently it cannot be applied to preferences for which a demand function does not exist, such as preferences which are only weakly convex or concave preferences. Therefore an extended definition is needed. Naturally an extended definition should coincide with the conventional definition for all preferences where the latter is applicable, namely for strongly convex preferences. Thus for strongly convex preferences all candidate definitions will necessarily be equivalent to each other and



(a) Concave preference (b) W-preference

Figure 3.2: Non-convex preferences

the conventional definition. This means that conventional definition cannot decide which of the candidate extended definitions is more useful or natural.

The tangency-characterization and the slope-characterization are the candidates for the extended definition of normality. If the preference is locally convex at x^0 then the characterizations are equivalent as previously argued, but it turns out that if the preference is locally strongly concave then they contradict each other. To illustrate this consider the strongly concave preference represented by $u(x_1, x_2) = (x_1)^2 + (x_2)^2$, shown in figure 3.2a, or the W-preference which at some points is locally concave, but at others locally convex, shown in 3.2b.

Points of tangency move to the "right" as we increase income, from a to b, so the tangency-characterization would classify good 1 as normal. By contrast the slope-characterization would classify good 1 as inferior since the slope of the indifference curve is flatter at c than at b. Neither of these classifications can be said to be incorrect at this point, it is a question of judgment which characteristic to pick as the defining one. Analytically the contradiction of the characterizations is

mirrored in the fact that the denominator in the tangency-condition, the bordered Hessian of u, is negative if the utility function is strictly quasi-convex.

In both these examples one might argue that a demand function almost exists, as demand is not unique only for very few prices, perhaps only for a measure zero set of prices. But the problem with defining normality as a local property is not existence or non-existence of the demand function but rather existence of the inverse demand function on the interior. No point at which the preference is locally concave is ever demanded, thus for the globally concave preference no interior point is ever demanded, while for the W-preference no point on the"Λ-part of the W" is ever demanded. Thus recourse to the conventional definition via the demand function cannot be taken for a local definition of normality which requires that the image of the demand function contains all interior x.

For three reasons I suggest that normality should be defined by employing the slope-characterization rather than the tangency-characterization: First the slope-characterization is easier to imagine and verify geometrically, second in its analytical form the slope characterization is simpler than the tangency-characterization as the former just consists of the numerator of the latter, and third, as will be shown later, defining normality by the slope-condition makes convexity an implication of normality in all goods.

***Definition* 3.2**

A preference is called normal in good 1 at x^0 *if* $\dfrac{u_1(x)}{u_2(x)}$ *is increasing in* x_2 *at* x^0.

***Fact* 3.1**

If u *is twice continuously differentiable then good 1 is normal at* x^0 *if and only if* $u_{12}(x^0)\,u_2(x^0) - u_{22}(x^0)\,u_1(x^0) \geq 0$.

Since a constant good is defined as a good that is both normal and inferior, a preference is constant in good 1 at x^0 iff $u_{12}(x^0)\,u_2(x^0) - u_{22}(x^0)\,u_1(x^0) = 0$. The sign of $u_{12}u_2 - u_{22}u_1$ is invariant to monotone transformations of the utility

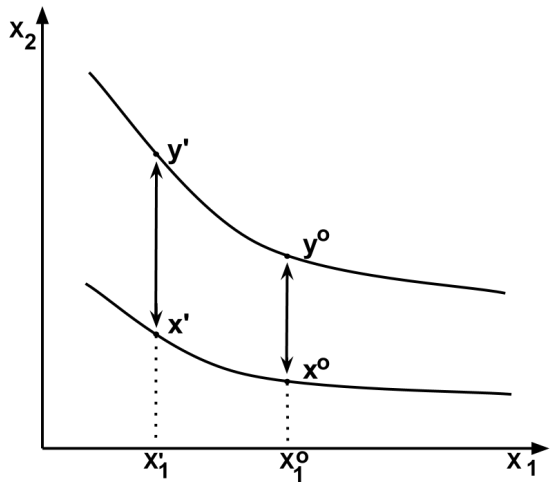


Figure 3.3: Distance between indifference curves.

function, and thus the definition indeed defines normality as property of the preference independently of the particular utility representation.

3.4 Defining normality on the preference itself

In choice theory the usual approach is to define properties directly on the preference rather than as a property of the utility, let alone the demand function. One could closely follow the previous definition, assume differentiability of the preference relation and consider how the direction of the normal vector of the indifference curve in x^0 changes as x_2 is increased. However, there is a simpler approach, that does not require differentiability, but merely continuity (and strong monotonicity) of the preference relation: The extended definition of normality implies that the"vertical" distance between two indifference curves declines as x_1 increases, as illustrated in figure 3.3.

Thereby for continuous, strongly monotone preference relations on $\mathbb{R}^2_{++}$ define, denoting $e_1 \equiv (1,0)$ and $e_2 \equiv (0,1)$:

***Definition* 3.3**

A preference is called normal in good 1 at x if there exists $\bar{\varepsilon} > 0$ such that for all y

with $x \sim y$, $\|x, y\| < \overline{\varepsilon}$, ε *with* $0 < \varepsilon < \overline{\varepsilon}$*:*

$x_1 \geq y_1$ *if and only if* $x + \varepsilon e_2 \succcurlyeq y + \varepsilon e_2$ *and* $y - \varepsilon e_2 \succcurlyeq x - \varepsilon e_2$.

Defining inferiority analogously implies that a preference is constant in good 1 at x iff there exists $\overline{\varepsilon} > 0$ such that for all y with $x \sim y$, $\|x, y\| < \overline{\varepsilon}$, and ε with $0 < \varepsilon < \overline{\varepsilon}$: $x + \varepsilon e_2 \sim y + \varepsilon e_2$ and $x - \varepsilon e_2 \sim y - \varepsilon e_2$. Recall that if good 1 is constant then good 2 is quasi-linear to see the connection between this local definition of a good as constant and Rubinstein's definition of a good as quasi-linear. In this paper's notation Rubinstein's definition for quasi-linearity of good 2 is:

***Definition* 3.4** (**Rubinstein's quasi-linearity**)

A preference is quasi-linear in good 2 if:

$x \succcurlyeq y$ *implies* $x + \varepsilon e_2 \succcurlyeq y + \varepsilon e_2$.

But that condition is equivalent to $x \sim y$ implies $x + \varepsilon e_2 \sim y + \varepsilon e_2$, which is the global equivalent of the above characterization of good 1 as constant.

3.5 Normality in both goods implies convexity

This section investigates the relationship between normality in the extended definition and convexity, and shows that if a preference relation is (locally) normal in all goods then it is (locally) convex.

***Proposition* 3.1**

If a preference relation is normal in all goods then it is convex.

Similarly if a preference relation is inferior in all goods then it must be concave. The two of them together imply that if a preference relation is constant in all goods, i.e. both weakly normal and weakly inferior in all goods, then the preference

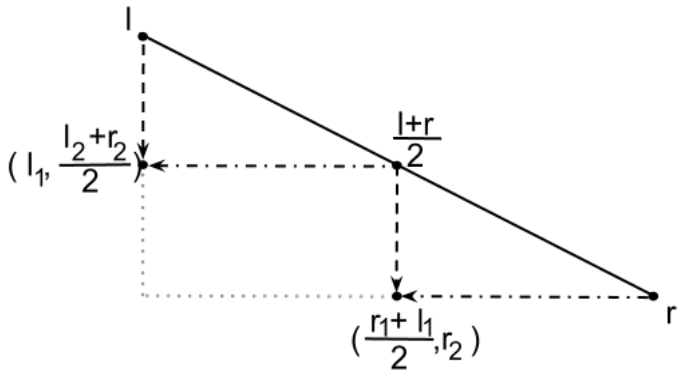


Figure 3.4: Normality implies convexity.

must be both weakly convex and weakly concave. Therefore its indifference curves are straight lines and we have the case of perfect substitutes.

Showing the proposition analytically is straightforward:

$$\begin{array}{rr} \text{Good 1 normal} \Rightarrow & u_{12}u_2 - u_{22}u_1 \geq 0 \mid \cdot u_1 \\ \text{Good 2 normal} \Rightarrow & +u_{12}u_2 - u_{22}u_1 \geq 0 \mid \cdot u_2 \\ \hline u \text{ quasi-concave} \Leftarrow & 2u_{12}u_1u_2 - u_{11}u_2 - u_{22}u_1 \geq 0 \end{array}$$

Showing the proposition directly with the preference relation might help the intuition more, especially since figure 3.4 accompanies the proof:

Proof. It suffices to show that for any two points l, r with $l \sim r$, and $l_1 < r_1$ (and thus $l_2 > r_2$): $\frac{l+r}{2} \succsim l$. Prove by contradiction so suppose $l \succ \frac{l+r}{2}$:
Since $r_1 > l_1$, normality in good 1 implies $l - \frac{l_2 - r_2}{2}e_2 \succ \frac{l+r}{2} - \frac{l_2 - r_2}{2}e_2$, i.e. $\left(l_1, \frac{l_2 + r_2}{2}\right) \succ \left(\frac{l_1 + r_1}{2}, r_2\right)$.
Since $r_2 > l_2$ (and $r \succ \frac{l+r}{2}$), normality in good 2 implies $r - \frac{r_1 - l_1}{2}e_1 \succ \frac{l+r}{2} - \frac{r_1 - l_1}{2}e_1$, i.e. $\left(\frac{l_1 + r_1}{2}, r_2\right) \succ \left(l_1, \frac{l_2 + r_2}{2}\right)$, a contradiction to the prior sentence. □

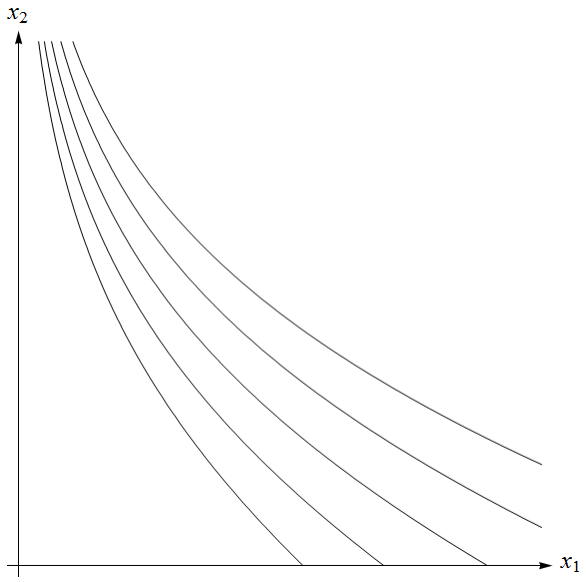


Figure 3.5: Modified Liebhafsky preferences.

3.5.1 Convexity does not imply that both goods are normal

The proposition stated that if all goods are normal then the preference is convex. The analogon of the proposition for inferior goods states that if all goods are inferior then the preference is concave. This implies that a convex preference must have at least one normal good. But do all goods have to be normal? It is well-known that the answer is no. Examples of well-behaved preferences that are strongly convex and have a (globally) inferior good seem to be missing in the literature. Therefore I will now construct an example of a preference that is both strongly convex and strongly inferior in one good. The utility function is a modification of a utility function proposed by Liebhafsky (1969). The modification makes the preference globally convex, which is not the case for the original Liebhafsky-preference.

***Example* 3.1**

The modified Liebhafsky preference is given by $u(x_1, x_2) = 4 \ln x_1 + \frac{1}{2}(x_2 + 3)^2$.

As $u_{12}u_2 - u_{22}u_1 = -\dfrac{4}{x_1} < 0$ good 1 is strongly inferior, while by $u_{12}u_1 - u_{11}u_2 = \dfrac{x_2 + 3}{x_1^2} > 0$ good 2 is strongly normal.

3.6 Conclusion

Starting with the usual definition of normality of a good as a global property of the demand function, this paper gives two equivalent characterizations of normality as a local property of a preference, the tangency- and the slope-characterizations. Geometrically the tangency-characterization states that the tangency point increases in good 1, while the slope-characterization states that the slope of the indifference curve through a point gets larger as x_2 is increased. To extend the definition of normality to weakly convex and even locally or globally concave preferences one of the two characterizations has to be chosen, as they contradict each other at points of local concavity. I choose the slope-characterization as it is the more intuitive one.

In chapter 4 it is shown that this new generalized definition has the elegant feature that (local) normality in all goods implies that the preference is (locally) convex. The converse is not true, a counterexample, the modified Liebhafsky preference shows that a strongly convex preference can still have an inferior, even globally inferior, good.

A good that is on the border between normal and inferior, i.e. a good that is both weakly normal and weakly inferior, I call constant. For a constant good demand is constant in income and the distance between two indifference curves is constant. It turns out that constant goods are the entourage of a quasilinear good. Constant goods allow for a novel perspective on quasi-linear preferences: In the case of two goods, if one good is constant then the other must be quasi-linear and vice versa. Thus the definition of a preference being locally constant in a good suggests a definition of quasi-linearity as a local property in a good. Closely related to this definition of local quasi-linearity of a preference is a definition by Rubinstein (2010) which is the global version of it.

Appendix

***Example* 3.2** (**Quasi-linear preference**)

$u(x_1, x_2) = \sqrt{x_1} + x_2$

Good 1 is constant, good 2 is normal and quasi-linear.

***Example* 3.3** (**Perfect substitutes**)

$u(x_1, x_2) = x_1 + x_2$

Both goods are constant (and also quasi-linear).

***Example* 3.4** (**CES**)

$u(x_1, x_2) = x_1^{\alpha} + x_2^{\alpha}$, $0 < \alpha < 1$.

Both goods are normal.

***Example* 3.5** (**Cobb-Douglas**)

$u(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, $0 < \alpha < 1$.

Both goods are normal.

***Example* 3.6** (**non-differentiable preference**)

$$u(x_1, x_2) = \begin{cases} x_1^{\frac{1}{3}} x_2^{\frac{2}{3}}, x_1 \leq x_2 \\ x_1^{\frac{2}{3}} x_2^{\frac{1}{3}}, x_1 > x_2 \end{cases}.$$

Both goods are normal.

References

Bergstrom, T., L. Blume, and H. Varian (1986). On the private provision of public goods. *Journal of Public Economics 29*, 25–49.

Besley, T. and S. Coate (1991). Public provision of private goods and the redistribution of income. *American Economic Review 81*(4), 979–84.

Bilancini, E. and L. Boncinelli (2010). Preferences and normal Goods: an easy-to-check necessary and sufficient condition. *Economics Letters 108*(1), 13–5.

Cummins, J. D. and O. Mahul (2004). The demand for insurance with an upper limit on coverage. *American Risk and Insurance Association 71*(2), 253–64.

Gertler, P., L. Locay, and W. Sanderson (1987). Are user fees regressive? The welfare implications of health care financing proposals in Peru. *Journal of Econometrics 36*, 67–88.

Green, J. and C. M. Kahn (1983). Wage-Employment contracts. *Quarterly Journal of Economics 98*, 173–87.

Kreps, D. M. (1988). *Notes on the Theory of Choice*. Boulder, CO: Westview Press.

Kreps, D. M. (1990). *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press.

Leroux, A. (1987). Preferences and normal goods: a sufficient condition. *Journal of Economic Theory 43*(1), 192–9.

Liebhafsky, H. (1969). New thoughts about inferior goods. *American Economic Review 59*(5), 931–4.

Nelson, W. R. (2001). Incorporating fairness into game theory and economics: comment. *American Economic Review 91*(4), 1180–3.

Rubinstein, A. (2011). Lecture notes in microeconomic theory. Revised version: http://arielrubinstein.tau.ac.il/Rubinstein2007.pdf.

Shitovitz, B. and M. Spiegel (2001). Stable provision vs. Cournot-Nash equilibria in pure public good economies. *Journal of Public Economic Theory 3*(2), 219–24.

Small, K. A. and H. Rosen (1981). Applied welfare economics with discrete choice models. *Econometrica 49*(1), 105–30.

Varian, H. R. (1992). *Microeconomic Analysis* (3d ed.). New York, NY: W.W. Norton.

Wilson, J. D. (1991). Optimal public good provision with limited lump-sum taxation. *American Economic Review 81*(1), 153–66.

Chapter 4

Transformation of a dimension

Abstract: This chapter considers order-preserving (i.e. monotone) transformations of a single *dimension* of a preference relation. The fact that order-preserving transformations of the *utility function* do not change the underlying preference is well-known. But order-preserving transformations of a *dimension* change the preference, a binary relation on a metric set, even while the actual choice behavior remains unchanged. For example a liquid good like milk may be measured in liters or gallons, the preference changes, but the transformation is ratio-preserving (i.e. linear) and so the dimension is called cardinal. Dimension that measure a quantity are naturally cardinal. By contrast dimensions that capture the quality of a service, the size of a physical good or the strength of a political preference are not naturally cardinal. For example, the size of a TV set with a fixed aspect ratio can be measured by the screen diagonal or the screen area. The information content is the same. But the transformation from one to another is merely order-preserving. Such changes in the measure change the economist's description of the choice behavior, the binary relation, but obviously do not alter the behavior itself. These transformations can change the properties of preferences. Economically relevant properties such as homotheticity, convexity and normality are invariant to linear transformations, but not to every ordinal transformation. This paper establishes that while homotheticity is invariant to linear transformations, convexity is invariant to convex transformations, and monotonicity is invariant to monotone transformations. The case of normality is a bit more complicated: normality in good

1 is invariant to monotone transformations of dimension 1, and invariant to convex transformations of dimension 2.

4.1 Introduction

This chapter asks whether properties of a preference relation like convexity and normality are invariant to order-preserving (monotone) but not necessarily ratio-preserving (linear) transformations of a dimension. For example dimension 1 might represent the screen diagonal of a certain make and model of TV that is available in different sizes. Given a fixed aspect ratio one could equally well measure the size of the TV by the screen area rather than the diagonal of the screen. Since area is a quadratic function of length, it is not true that the preference that represents behavior using one measure is convex, homothetic or normal if and only if the preference that represents the same behavior in the other measure is. It would, however, be true if the transformation was linear, an intuitive point this chapter formalizes: convexity, normality and homotheticity are all invariant to linear transformations of a dimension. Quantities have a natural linearity, so the problem this chapter discusses does not arise in models that only include dimensions that give a quantity of a good.

Transformations of a dimension are distinct from the well-known transformations of the utility function itself. In that case any positive monotone transformation of a utility function still represents exactly the same preference relation. That result does not apply to transformations of a dimension. Indeed the problems considered in this paper exist even if one uses preferences only, and makes no use of utility functions.

With the transformation of a dimension, the preference, a binary relation on an Euclidean space, *itself* changes. In other words the indifference curves change. The behavior represented does not change, but the scale of dimension 1 and thus the interpretation of each point in the space changes.

Figure 4.1 illustrates a central result of this chapter: The black line is a single indifference curve of a convex preference. The grey curve to the right represents the same set of physical bundles after a convex transformation of dimension 1. The

grey curve to the left does the same exercise for a concave transformation of dimension 1. Observe that in this example convexity is invariant to the convex transformation, but not to the concave one. This chapter develops this observation into a more general statement: convexity of a preference is always preserved under a convex transformation, but may or may not survive non-convex transformations. Moreover, for any non-convex transformation there exists a convex preference that loses its convexity under the transformation.

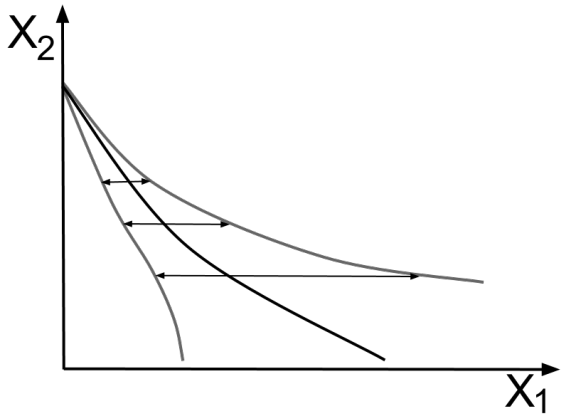


Figure 4.1: Convex preferences are invariant to convex transformations.

Besides convexity I investigate the invariance of homotheticity and normality. Homotheticity turns out to be invariant to all linear transformations, normality in good 2 is invariant to all convex transformations (of dimension 1), and normality in good 1 is invariant to all monotone transformations (of dimension 1). Finally monotonicity is invariant to all monotone transformations. Note that the invariance of normality in good 1 to monotone transformations of dimension 1 employs the definition of normality from chapter 3 which extends the conventional definition to non-convex preferences.

To the best of my knowledge the problem of invariance of properties of preference relations to non-linear transformations of a dimension has so far not been treated in the literature, other than in passing. The reason for this is presumably that economics has traditionally considered each dimension as representing the *quantity* of

a homogenous goods such as wool or wine. Quantities are additive and therefore transformations one would want to consider naturally limit themselves to ratio-preserving, i.e. linear ones. For discrete (indivisible) goods such as apples there is even a natural unit. Thus for discrete homogeneous goods there is no reason to even worry about linear transformations, while for non-discrete homogeneous goods such as milk or sugar the relevant transformations such as converting from gallons to liters or pounds to kilograms are always linear. The dimensions economists historically were interested in were cardinal. And it happens to be the case that properties of binary relations of economic interest are invariant to linear transformations. Yet if something other than quantity is measured, say size or quality, there is no reason why the dimension would be cardinal.

Moussa and Rosen (1978) investigate the optimal menu of price and quality of a monopolist. In their setting consumer preferences are perfect substitutes over two goods, a numeraire and quality. They point out that with an appropriate "redefinition of the units in which 'quality' is measured" their treatment generalizes to preferences that are quasi-linear, where the numeraire is the quasilinear good and quality the constant good. While Moussa and Rosen assume preferences are perfect substitutes and let cost be an increasing convex function of quality, Maskin and Riley (1984), building on Moussa and Rosen, assume that cost is linear in quality but preferences are quasi-linear. Both articles indirectly make use of the fact that a convex transformation preserves convexity and a fortiori normality and constancy.

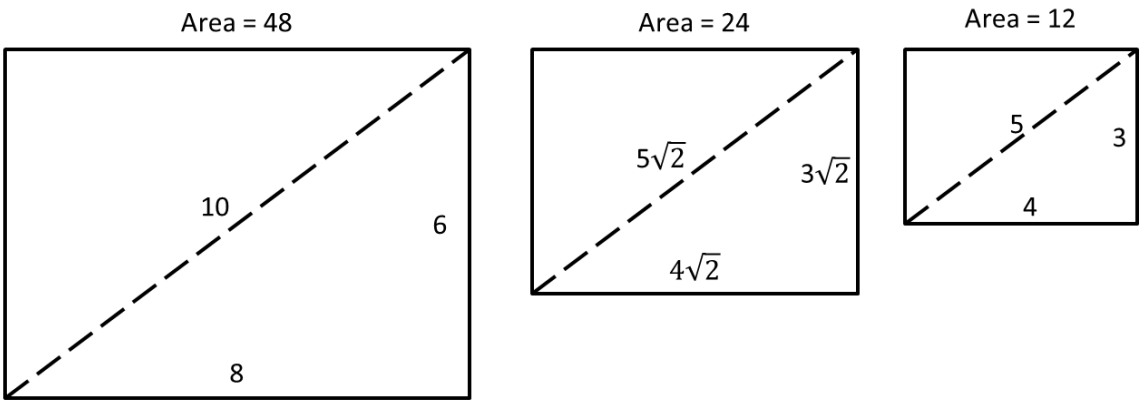
Ordinal dimensions have become more important with rising real incomes as the share of consumption of goods where quality rather than quantity is the main source of variation in value has risen. Examples of such goods include electronic devices, cars, homes, household appliances, education, medical procedures, and restaurant meals.

Transformations of a dimension being a novel topic, section 2 discusses the research question with the example of the size of a TV set. Size is not the most relevant empirical application, but allows us to ask the research question in particular simplicity. Thereafter section 3 develops the framework of an observable preference and its potential representations in Euclidean spaces, and thereby fixes ideas and gives the groundwork definitions. Section 4 first points out the invariance of properties of binary relations that apply even to non-metric sets (e.g. completeness), and then investigates those that apply to Euclidean sets (e.g. convexity) and in turn shows which ones are invariant to linear, convex and monotone transformations. Section 5 concludes.

4.2 Example: Screen area and diagonal

A consumer has preferences over two dimensions. Dimension 2 is a standard, cardinal dimension that gives the quantity of a numeraire good. Dimension 1 is merely ordinal and gives the size of a particular make and model of a TV set. The quantity of TV sets purchased is fixed and equals 1. Regardless of its size the TV set has a fixed aspect ratio of width:height = 4:3.

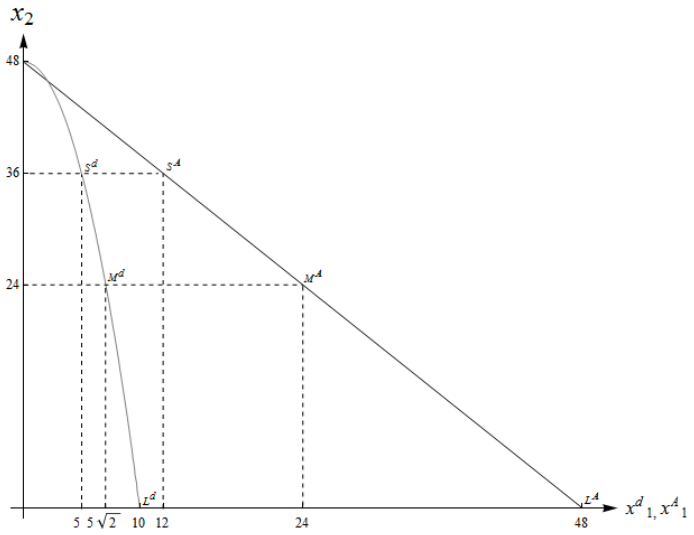
Figure 4.2a shows three such TV sets, a small (S), medium (M) and large TV set (L). In terms of area the size of the TV just doubles from small to medium and from medium to large. The screen diagonal however is just $\sqrt{2}$-times bigger for the medium versus the small set, and the large versus the medium set. Given the constant aspect ratio both measures, area and diagonal, convey exactly the same information. It is not clear that either measure would be superior to, or more natural than the other. Suppose the consumer reveals her preference $\succcurlyeq$ to be such that she is indifferent between the bundle of the small TV and 36 units of the numeraire, and the bundle of the medium TV and 24 units of the numeraire, i.e. $(S, 36) \sim (M, 24)$,



(a) TV sets with aspect ratio 4:3.

No metric	Area A	Diagonal d
(S,36)	(12,36)	(5,36)
(M,24)	(24,24)	(5$\sqrt{2}$,24)
(L,0)	(48,0)	(10,0)

(b) Bundles of TV set and numeraire.



(c) Interpreting axis 1 as screen area or diagonal.

and that she is also indifferent between the bundle of the medium TV and 24 units of the numeraire and the large TV and 0 units of the numeraire, i.e. $(M, 24) \sim (L, 0)$. It is possible to ask whether $\succcurlyeq$ is complete or transitive. For example the former entails asking whether at least one of $(S, 36) \succcurlyeq (L, 0)$ or $(L, 0) \succcurlyeq (S, 36)$ holds, while the latter entails asking whether $(S, 36) \sim (L, 0)$. Asking whether $\succcurlyeq$ is convex is however an ill-posed question, as it is not clear for example whether $(M, 24)$ is a convex combination of $(S, 36)$ and $(L, 0)$. While $24 = \frac{2}{3}36 + \frac{1}{3}0$, no such relationship holds for S, L and M. The concept of a convex combination of L and S is not defined. To consider convexity it is not enough that the preference $\succcurlyeq$ be a binary relation -it is-, it needs to be a binary relation on a *metric* set.

To define the preference as a binary relation on $\mathbb{R}^2_+$ the size of the TV needs to be measured. One economist observer chooses to measure the size of the TV set by screen area, and interprets the preference to be perfect substitutes between screen area and the numeraire. In figure 4.2c the straight bold line gives a single indifference curve of that preference. A utility representation of that preference is for example $u^A\left(x_1^A, x_2\right) = x_1^A + x_2$, where the superscript A indicates that screen size is measured by area. Denote this preference by $\succcurlyeq^A$. Another economist observer however measures size by the diagonal, and writes down the utility function $u^d\left(x_1^d, x_2\right) = \frac{12}{25}\left(x_1^d\right)^2 + x_2$ representing the preference relation $\succcurlyeq^d$. In figure 4.2c a single indifference curve of this concave preference relation is shown. Both $\succcurlyeq^A$ and $\succcurlyeq^d$ rank the three bundles as indifferent: $(12, 36) \sim^A (24, 24) \sim^A (48, 0)$, respectively $(5, 36) \sim^d \left(5\sqrt{2}, 24\right) \sim^d (10, 0)$. Indeed the rankings of $\succcurlyeq^A$ and $\succcurlyeq^d$ for any physical bundles are identical. Thus $\succcurlyeq^A$ and $\succcurlyeq^d$ are observationally equivalent. To verify this note that the screen diagonal x_1^d can be expressed as a function t of the diagonal: $x_1^d\left(x_1^A\right) = t\left(x_1^A\right) = 5\sqrt{\frac{x_1^A}{3 \cdot 4}}$, and therefore:

$$u^d\left(x_1^d\left(x_1^A\right), x_2\right) = \frac{12}{25}\left(x_1^d\left(x_1^A\right)\right)^2 + x_2 = x_1^A + x_2 = u^A\left(x_1^A, x_2\right).$$

The transformation t is strictly increasing and thus invertible, so the above logic works for the reverse as well. The utility functions u^A and u^d are *not* transformations of each other as there exists no function $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f\left(u^A\right) = u^d$. The well-known point that any positive monotone transformation of a utility function is again a utility function for the same preference does hold for each of u^A and u^d separately. If f is a positive monotone function $\mathbb{R} \longrightarrow \mathbb{R}$ then u^A and $f\left(u^A\right)$ are both utility representations of $\succcurlyeq^A$. Yet neither u^A nor $f\left(u^A\right)$ represent $\succcurlyeq^d$. Indeed they cannot since $\succcurlyeq^A$ and $\succcurlyeq^d$ are not the same preference relation. Yet with the respective

interpretations of dimension 1, interpretations of what a point in Euclidean space means in the observable world, both $\succcurlyeq^A$ and $\succcurlyeq^d$ represent the same choice behavior.

Observe that $\succcurlyeq^d$ is a concave preference, yet $\succcurlyeq^A$ is convex (and concave). Thus convexity is one example of a property of a preference relation that is not invariant to monotone transformations of a dimension.

4.3 Revealed preference and its representations

Choice theory understands a preference to be a binary relation that is inferred from observable behavior. The consumer chooses one of the affordable bundles of goods. The economist observer may have beliefs about the consumer's mental process including whether and in what units the consumer thinks. The observer may believe that a shopper purchasing milk in the US thinks in gallons, and that a shopper purchasing milk in France thinks in liters, but what is observed is the behavior and the situation, what the consumer purchases given what budget. Indeed the consumer needs not even be literate or numerate to make choices, and therefore for her behavior to be describable by a preference. Even stone age humans presumably had preferences. Further afield economists have described animal behavior in terms of preferences without claiming that these animals could count or measure (e.g. Battalio et al. 1985). To summarize, at least in principle, a preference relation is inferred from behavior, from a choice between various physical bundles of goods, the standard approach of revealed preference in economics since at least Samuelson (1948).

The inferred or observable preference is therefore a binary relation on the *set of all bundles* under consideration. The set of bundles by itself has no metric. This preference is still a binary relation, but it is a binary relation on a set without an order let alone a metric. On such sets certain properties of binary relations, for example completeness, symmetry and transitivity are already well-defined. But

other properties of binary relations are not well-defined on such sets, for example monotonicity, convexity, normality, and homotheticity.

Let $\succcurlyeq^*$ denote the observable preference as revealed by choice behavior between bundles of physical goods. Denote the set of all bundles by S. In most empirical applications all dimensions but one will be cardinal, and only one dimension will be merely ordinal. For ease of exposition I focus on the two-dimensional case, with dimension 1 being merely ordinal and dimension 2 being cardinal.

The observer can choose to measure each good in the bundles by a metric of her choice. With that metric she can write down a preference relation $\succcurlyeq^A$ which is a binary relation on some set $A = A_1 \times A_2$. We assume that for dimensions $k = 1, 2$ we have $A_k = [0, \overline{a}_k]$ with $\overline{a}_k > 0$ or $A_k = \mathbb{R}_0^+$.

Good 1 is an indivisible good of which the consumer consumes exactly one unit which has a one-dimensional attribute, for example quality or size. Dimension 1 is ordinal as it gives this quality or size. Good 2 is a homogenous good, so dimension 2 is ordinal. Denote a typical element of A by $\left(x_1^A, x_2^A\right)$.To be able to do so without having to throw away information the following assumption should hold:

***Assumption* 4.1**

There exist bijective functions $m_1^A : S_1 \to A_1$, $m_2 : S_2 \to A_2$.

The preference relation $\succcurlyeq^A$ and an observable preference $\succcurlyeq^*$ are different representations of the same reality if they order bundles that are mapped into each other in the same way, therefore we say:

***Definition* 4.1**

A preference $\succcurlyeq^A$ represents $\succcurlyeq^*$*on* A *if for all* s, t *in* S*:* $s \succcurlyeq^* t$ *if and only if* $\left(m_1^A(s_1), m_2^A(s_2)\right) \succcurlyeq^A \left(m_1^A(t_1), m_2^A(t_2)\right)$

We focus attention on dimension 1 only, and therefore make the assumption that dimension 2 is identical for any space that the economist observer chooses:

***Assumption* 4.2**

$m_2^A = m_2^B$

In the following we can therefore drop the superscripts and simply write $m_2 (\equiv m_2^A \equiv m_2^B)$ and $x_2 (\equiv x_2^A \equiv x_2^B)$. We restrict attention to cases where there is ordinal agreement between the measures:

***Assumption* 4.3** (**ordinal agreement**)

For all A, B: $x_1^A \geq x_1^{A\prime}$ if and only if $x_1^B \geq x_1^{B\prime}$

***Fact* 4.1**

For all A, B there exists a strictly monotone transformation $t : A_1 \to B_1$ with $t(0) = 0$.

To see that such a t exists, let's construct it by $t\left(x_1^A\right) = m_1^B\left(m_1^{A-1}\left(x_1^A\right)\right)$. This implies that t has an inverse function which we shall denote by t^{-1}. Note that $t^{-1}(0) = 0$, therefore there exists a natural minimum of dimension 1 and all measures assign the number 0 to it. For example temperature has a natural minimum, 0 Kelvin, so we would not allow measures like Fahrenheit or Celsius that assign negative numbers to the absolute minimum temperature.

4.4 Invariance

***Fact* 4.2** (**completeness**)

$\succcurlyeq^*$ is complete if and only if $\succcurlyeq^A$ is complete.

***Fact* 4.3** (**transitivity**)

$\succcurlyeq^*$ is transitive if and only if $\succcurlyeq^A$ is transitive.

Analogous statements can be made for some other properties of binary relations that do not require a metric such as acyclicity, symmetry etc. However, we cannot make analogous statements for properties like monotonicity, convexity, normality or homotheticity, as they are well defined only for $\succcurlyeq^A$, not for $\succcurlyeq^*$.

Euclidean properties like convexity are not defined for the observable preference $\succcurlyeq^*$, but only for its representations on Euclidean spaces. Therefore we now investigate under what circumstances a property of a preference $\succcurlyeq^A$ which represents $\succcurlyeq^*$ on A, holds for $\succcurlyeq^B$ which represents $\succcurlyeq^*$ on B.

We focus on three main properties of binary relations which are defined on Euclidean spaces and commonly used in economics: monotonicity, normality and convexity. For transformations of dimension 1 we will consider three narrowing classes of transformations: positive monotone transformations, convex transformations and linear transformations. For simplicity we focus on preferences that are representable by a twice continuously differentiable utility function and on transformations that are twice continuously differentiable.

***Assumption* 4.4** (**monotonicity**)

For all A, x^A:

$\left(x_1^A, x_2\right) \succcurlyeq^A \left(x_1^{A\prime}, x_2\right)$ *if and only if* $x_1^A \geq x_1^{A\prime}$, *and*

$\left(x_1^A, x_2\right) \succcurlyeq^A \left(x_1^A, x_2'\right)$ *if and only if* $x_2 \geq x_2'$.

***Assumption* 4.5** (**differentiability**)

There exists a twice continuously differentiable utility representation u^A of $\succcurlyeq^A$. Any transformation t is twice continuously differentiable and satisfies $t_1 > 0$.

***Fact* 4.4**

If $\succcurlyeq^A$ represents $\succcurlyeq^*$ on A, u^* represents $\succcurlyeq^*$, and u^A represents $\succcurlyeq^A$, then for all s, t in S: $u^*(s) \geq u^*(t)$ if and only if $u^A\left(m_1^A(s_1), m_2(s_2)\right) \geq u^A\left(m_1^A(t_1), m_2(t_2)\right)$.

Now let us define a utility function u^B on B and then argue that it indeed represents $\succcurlyeq^*$ on B

$u^B\left(x_1^B, x_2\right) \equiv u^A\left(t^{-1}\left(x_1^B\right), x_2\right)$, where t^{-1} is the inverse function which exists by strict monotonicity of t.

4.4.1 Linear transformations

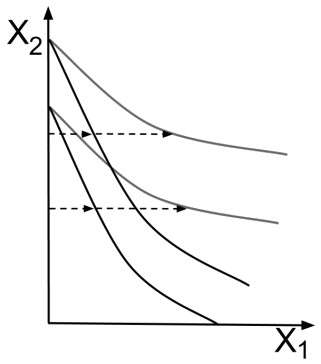


Figure 4.2: Linear transformation of axis 1.

***Proposition* 4.1** (Homotheticity is invariant to linear transformations)

Homotheticity of a preference is invariant to any positive linear transformation t : $t(x_1) = \alpha x_1$, $\alpha > 0$.

Figure 4.2 illustrates a linear transformation which is a stretch parallel to the x_1 axis from the x_2-axis, here by approximately factor 2. The black indifference curves are the original ones corresponding to $\succcurlyeq^A$, the grey ones correspond to $\succcurlyeq^B$.

4.4.2 Convex transformations

Convexity of a preference is invariant to convex transformations

***Proposition* 4.2**

Convexity of a preference is invariant to convex transformations of a dimension.

***Claim* 4.1**

If for a transformation t *the convexity of all convex preferences is invariant to* t *then the* t *is convex.*

A convex transformation of a dimension just says that the amount of stretching at a certain x_1 is increasing in x_1.

Now what the above two results do not preclude is that a particular preference is invariant to a particular non-convex transformation. As an example consider the Cobb-Douglas preference relation represented by $u^A\left(x_1^A, x_2\right) = x_1^A x_2$, which of course is a convex preference relation, and consider the concave transformation $t\left(x_1^A\right) = \sqrt{x_1^A}$. Then one utility representation on B is $u^B\left(x_1^B, x_2\right) = \left(x_1^B\right)^2 x_2$. But this utility function again represents a convex preference, so even after we apply a concave transformation the preference remains convex. Thus it can happen that a particular concave transformation does not destroy convexity of a particular preference. Yet for any particular preference there exists a transformation that does.

***Fact* 4.5**

For any convex preference there exists a monotone transformation that does not preserve convexity.

For the symmetric Cobb-Douglas preference in the previous example we showed that applying the concave transformation of taking the square root does not destroy convexity, so now let's give a transformation for that preference that does destroy convexity.

***Example* 4.1 (convexity not invariant to concave transformation)**

$u^A\left(x_1^A, x_2\right) = x_1^A x_2$, $t\left(x_1^A\right) = 1 - e^{-x_1^A}$. *Note that if* $A_1 = \mathbb{R}_+$ *then* $C_1 = (0,1)$. *The inverse of the transformation is* $t^{-1}\left(x_1^C\right) = -\ln\left(1 - x_1^C\right)$ *and so one representation of the preference relation is* $u^C\left(x_1^C, x_2\right) = \ln\left(\frac{1}{1-x_1^C}\right) \cdot x_2$, *which is not quasi-concave globally so the preference is not convex.*[1]

[1]To verify this note that the derivatives are $u_1 = \frac{1}{1-x_1^C} \cdot x_2 + \ln\left(\frac{1}{1-x_1^C}\right)$, $u_2 = \ln\left(\frac{1}{1-x_1^C}\right)$, $u_{11} = -\frac{1}{\left(1-x_1^C\right)^2} \cdot x_2 + \frac{1}{1-x_1^C}$, $u_{12} = \frac{1}{1-x_1^C}$ and $u_{22} = 0$. Note $0 < x_1^C < 1$. But then $\ln\left(\frac{1}{1-x_1^C}\right) \cdot \frac{1}{1-x_1^C}\left(1 - \frac{1}{\left(1-x_1^C\right)} \cdot x_2\right)\left[2\frac{1}{1-x_1^C} - \ln\left(\frac{1}{1-x_1^C}\right)\right]$ can be negative.

Normality in good 2 is invariant to convex transformations

***Definition* 4.2** (**normality**)

We say that a preference relation on $\mathbb{R}^2_+$ represented by u is normal in good i if $u_{ij}u_j - u_{jj}u_i \geq 0$.

***Proposition* 4.3** (**good 2 normal invariant to convex transformations of dimension 1**)

Normality in good 2 of a preference is invariant to any convex transformation t of dimension 1.

4.4.3 Monotone transformations

***Proposition* 4.4** (**good 1 normal invariant to monotone transformations of dimension 1**)

Normality in good 1 is invariant to any transformation t of good 1.

Figure 4.3 illustrates why this must be the case. Recall that normality of good 1 means that the indifference curves get steeper along parallels to the x_2-axis. No combination of stretchings and shrinkings of different parts of axis 1 can possibly alter the fact that the slope gets steeper along the dashed line.

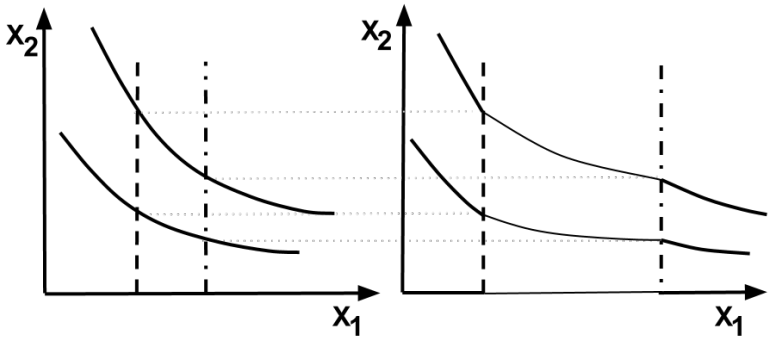


Figure 4.3: Monotone transformation of axis 1.

4.5 Conclusion

The present chapter examines a model with two goods, further research could extend this to three or more goods. To what extent the current results might be related to results in Milgrom and Shannon (1994)'s work on monotone comparative statics should be investigated. Traditionally comparative statics requires making assumptions of cardinal properties such as convexity or linearity of feasible or desirable sets, therefore they develop "methods for comparative statics analysis using only conditions that are ordinal" (ibid., p.158). In a sense chapters 3 and 4 might be in line with their approach: normality, which is conventionally defined as a cardinal condition is defined more generally in chapter 3. The more general definition is then shown to be in chapter 4 to be an ordinal condition.

The question in this chapter arises when a dimension cannot be considered naturally cardinal. Cardinality of a dimension is natural to assume for quantity, but not when the dimension measures the quality of a good or service. For example chapter 1 assumes normality of good 1, the quality of a medical service, so it is natural to ask whether that normality depends on the particular measure of quality being chosen. It turns out that while an order-preserving transformation of dimension 1 may result in the loss of convexity, it cannot result in the loss of normality of good 1, as long as normality is defined with the extended definition of chapter 3. This chapter investigated such invariance results of preferences to order-preserving transformations, the table summarizes the results:

	t strictly monotone	t convex	t linear
monotone	invariant	invariant	invariant
normal in good 1	invariant	invariant	invariant
normal in good 2	-	invariant	invariant
convex	-	invariant	invariant
homothetic	-	-	invariant

Homotheticity is invariant to linear transformations, but not strongly convex or concave ones, and consequently a fragile property. It seems that relying on homotheticity in theoretical or empirical work that involves an ordinal dimension raises concerns about the meaning of the assumption. Convexity is slightly less fragile. It is invariant to convex transformations. Normality in good 1 is invariant to order-preserving transformations of dimension 1. The same holds for monotonicity in good 1. The difference is that monotonicity in *good 1*, unlike normality in *good 1*, is invariant to order-preserving transformations of *dimension 2*. Normality in good 1 is invariant to convex transformations of dimension 2. Thus this chapter suggests that the assumption of normality of good 1 made in chapter 1 is meaningful even though the dimension of good 1, the quality of a medical service, is merely ordinal.

Appendix

It is useful to express the first and second derivatives of u^B in terms of u^A, t and their derivatives. For brevity we omit the arguments, which are $\left(x_1^B, x_2\right)$ for the derivatives of u^B, $\left(t^{-1}\left(x_1^B\right), x_2\right) = \left(x_1^A, x_2\right)$ for the derivatives of u^A, and $x_1^A = t^{-1}\left(x_1^B\right)$ for the derivatives of t.

$$u_1^B = \frac{u_1^A}{t_1}, u_2^B = u_2^A,$$
$$u_{11}^B = \frac{u_{11}^A - u_1^A \frac{t_{11}}{t_1}}{\left(t_1\right)^2}, u_{12}^B\left(x_1^B, x_2\right) = \frac{u_{12}^A}{t_1}, u_{22}^B = u_{22}^A.$$

Proof that homotheticity is invariant to linear transformations:

(Proof). If u is homothetic then for all $\lambda > 0$: $\frac{u_1(\lambda x)}{u_2(\lambda x)} = \frac{u_1(x)}{u_2(x)}$. Now $\frac{u_1^B(\lambda x)}{u_2^B(\lambda x)} = \frac{\frac{u_1^A(\lambda x)}{t_1(\lambda x_1)}}{u_2^A(\lambda x)} = \frac{1}{\alpha}\frac{u_1^A(\lambda x)}{u_2^A(\lambda x)} = \frac{1}{\alpha}\frac{u_1^A(x)}{u_2^A(x)} = \frac{1}{\alpha}\frac{\frac{u_1^B(x)}{t_1^{-1}(x_1)}}{u_2^B(x)} = \frac{u_1^B(x)}{u_2^B(x)}$. □

Proof that convexity is invariant to convex transformations:

(Proof). If $\succcurlyeq^A$ is convex then its utility representation u^A satisfies the bordered Hessian condition, that is $u_{11}^A \leq 0$, $u_{22}^A \geq 0$ and $2u_{12}^A u_1^A u_2^A - u_{11}^A\left(u_2^A\right)^2 - u_{22}^A\left(u^A\right)^2 \geq 0$. We show that the bordered Hessian condition holds for u^B as well. Note that by convexity of t we have $t_{11} \geq 0$.

1. $u_{22}^B = u_{22}^A \leq 0$.
2. $u_{11}^B = \frac{u_{11}^A - u_1^A \frac{t_{11}}{t_1}}{\left(t_1\right)^2} \leq 0$
3. $2u_{12}^B u_1^B u_2^B - u_{11}^B\left(u_2^B\right)^2 - u_{22}^B\left(u_1^B\right)^2$
   $= \frac{2u_{12}^A u_1^A u_2^A - u_{11}^A\left(u_2^2\right)^2 - u_{22}^A u_1^A + \frac{t_{11}}{t_1}u_1^A\left(u_2^A\right)^2}{\left(t_1\right)^2} \geq 0.$

□

Proof that normality in good 2 is invariant to convex transformations:

(Proof). $\succcurlyeq^A$ is normal in good 2, if and only if its utility representation u^A satisfies $u_{12}^A u_1^A - u_{11}^A u_2^A \geq 0$. We show the equivalent for u^B: $u_{12}^B u_1^B - u_{11}^B u_2^B = \frac{u_{12}^A}{t_1}\frac{u_1^A}{t_1} - \frac{u_{11}^A - u_1^A \frac{t_{11}}{t_1}}{(t_1)^2} u_2^A \geq \frac{1}{(t_1)^2}\left(u_{12}^A u_1^A - u_{11}^A u_2^A\right) \geq 0$. □

Proof that if all convex preferences are invariant to a transformation t, then t is convex:

(Proof). Consider the preference represented by $u^A\left(x_1^A, x_2\right) = x_1^A + x_2$, it is weakly convex. On B one of its utility representations is $u^B\left(x_1^B, x_2\right) = t^{-1}\left(x_1^B\right) + x_2$. Convexity of $\succcurlyeq^B$ would require that for all $\left(x_1^B, x_2\right)$ we have $2u_{12}^B u_1^B u_2^B - u_{11}^B\left(u_2^B\right)^2 - u_{22}^B\left(u_1^B\right)^2 \geq 0$. But $2u_{12}^B u_1^B u_2^B - u_{11}^B\left(u_2^B\right)^2 - u_{22}^B\left(u_1^B\right)^2 =$
$\frac{2u_{12}^A u_1^A u_2^A - u_{11}^A\left(u_2^2\right)^2 - u_{22}^A u_1^A + \frac{t_{11}}{t_1} u_1^A\left(u_2^A\right)^2}{(t_1)^2}$
$= \frac{t_{11}}{(t_1)^3}$. But since t is not convex there exists at least one x_1^A such that $t_{11}\left(x_1^A\right) < 0$. □

Proof that normality in good 1 invariant to any monotone transformation:

(Proof). Suppose a preference is normal in good 1 on A, then its utility representation u on A satisfies $u_{12}^A u_2^A - u_{22}^A u_1^A \geq 0$. We want to show that $u_{12}^B u_2^B - u_{22}^B u_1^B \geq 0$. Note that $u_{12}^B u_2^B - u_{22}^B u_1^B = \frac{u_{12}^A}{t_1} u_2^B - u_{22}^A \frac{u_1^A}{t_1} = \frac{1}{t_1}\left(u_{12}^A u_2^B - u_{22}^A u_1^A\right) \geq 0$. □

References

Battalio, R. C., J. H. Kagel, and D. N. MacDonald (1985). Animals' choices over uncertain outcomes: Some initial experimental results. *American Economic Review 75*(4), 597–613.

Maskin, E. and J. Riley (1984). Monopoly with incomplete information. *Rand Journal of Economics 15*(2), 171–196.

Milgrom, P. and C. Shannon (1994). Monotone Comparative Statics. *Econometrica 62*(1), 157–180.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory 18*, 301–317.

Samuelson, P. A. (1948). Consumption theory in terms of revealed preference. *Econometrica 15*(60), 243–53.