Essays in Applied Microeconomics

Péter Isztin

Institute of Economics

Supervisor:

Barna Bakó, PhD

Institute of Economics, Corvinus University of Budapest

Corvinus University of Budapest

Doctoral School of Economics, Business and Informatics

Essays in Applied Microeconomics

Ph.D Dissertation

Péter Isztin

Budapest, 2021

To the memories of my mother and Gary Becker

Contents

Introduction

In the following we present five essays that constitute the core of my PhD thesis. The essays connect with each other not so much in terms of one common overarching theme, but rather in terms of the methods of attacking the problems analyzed. Out of the five essays, two (Chapter 1 and 4) were written with Barna Bakó, one (Chapter 3) with Barnac, Zombor Berezvai and Enikő Vigh, and two essays (Chapter 3 and 5) are self-authored. The thesis will be structured as follows: in the introduction, we will go through common themes or methodological insights that connect the essays. In particular, we will outline a case for the continued relevance of applied theory in economic research and briefly present the results we arrived at in each of the essays. The remainder of the thesis consists of the essays themselves.

1 The place of applied theory within economic research

Intensive specialization is a conspicuous feature of modern economic research. In particular, there is an increasing gap between theory and empirical research. One strain of economic analysis that aims to bridge purely theoretical and purely applied or empirical economics is applied theory. Applied theory consists of the application of established economic theory to various social issues and problems. Various Nobel-prize winners, including Gary Becker or George Akerlof largely belong to this tradition, and yet the filed might be regarded as somewhat "old-fashioned". While in many parts of economic research, such as industrial organization, there are still a lot of applied theory papers, with practical insights and hypotheses which are then tested either within the same paper or in purely empirical studies, in some fields of economics, research has moved to a more purely empirical direction.This is particularly true in areas such as labor economics, where the role of theory has diminished a great deal in the past few decades. Meanwhile, (pure) economic theory is often highly abstract and lack direct empirical or policy relevance. Our view is that the best empirical papers either build on theoretical insights, or even have a theoretical part in them, and likewise, the best kind of theoretical papers build on empirical insights and aim to answer practical questions. In this thesis, we set out to contribute to both types

of research: most of the chapters are in the applied theory tradition, while one chapter is empirical, with a dose of theoretical intuition to underpin it.

A common theme across the chapters is that economics is defined not by its subject matter but by its method, or way of looking at the world. Among else, in the thesis we aim to shed light on the demand for information in product markets, persuasion involving supernatural claims in the context of the family and politics, the role of expected rents in human capital investment decisions, substitution and complementarity across different modes of transportation, and investing by parents in their children's character traits to influence their life outcomes. Throughout the chapters, decision-makers are assumed to be "rational" in terms of doing as best as they can in achieving their goals, as well as forward-looking. We use some insights from behavioral economics, such as the idea that for consumers information is not always a good, but we put these ideas in a standard neoclassical maximizing framework and importantly, in a market context. We also build on ideas that what we observe as preferences may be changed through conscious effort or investments (for example by parents), although it is important to stress that these kinds of investments can be fruitfully analyzed only if we assume stable "meta-preferences". It is our fervent hope that the presented essays, some of which have already been published, while others could be the basis of future publications, can shed light on the issues analyzed and can serve as inspirations for future theoretical and empirical research.

2 Overarching themes

Although the chapters may not be closely connected to each other, a number of themes connect them in some ways. Information and persuasion is a common theme of the first and the last two chapters.

Human development and investing in children within a family environment is another theme that runs through multiple chapters. We discuss family decision making about children in the chapter dealing with different kinds of human capital investment (Chapter 3), in the chapter dealing with supernatural persuasion (Chapter 4), and also in the last chapter which is about instilling values in children (Chapter 5). The economics of crime is also a theme that appears in more than one chapter. We discuss, among else, parents' as well as political leaders' actions aimed at discouraging criminal behavior among their offspring and their citizens.

Some of our themes build on what is widely characterized as "economic imperialism". Economic imperialism (see a useful overview by Lazear (2000)) is the application of basic economic (theoretical and empirical) tools to a wide range of social phenomena, including the family, politics, addictions and even religion. This thesis in particular deals with

the family: human capital, preference-formation and belief-formation within the family, religion and the use of religion to influence beliefs or preferences and the idea of political markets. When applying economic theory to areas other than strictly economic ones we obviously need to familiarize ourselves with preexisting ideas in those fields, such as sociology, psychology or evolutionary science. Nonetheless, it is obvious to us that the presented essays do not capture fully the underlying sociological and other insights as we build models that necessarily present a simplified version of reality. The economic approach is characterized by postulating maximizing behavior, the notion of equilibrium and also on stable preferences (Becker, 1992). Apart from these, Lazear also points to the notion of efficiency. In one chapter we seem to build a model of endogenous preferences, but, like Becker and Murphy (1988) we do so while maintaining the notion of stable meta-preferences, where the meta-preferences are either the preferences of the children or the parents shaping their children's preferences. Efficiency also plays a key role in some of the essays: certain social institutions such as the sexual division of labor, parental "indoctrination" and the use of religion in the public sphere are shown to serve an efficient "function": however, social institutions that were once efficient may not be efficient in a modern society. We emphasize that adaptions to changing circumstances do usually happen, and in that societies are continuously converging toward efficient outcomes. In this regard, the essays have a "functionalist" taste.

3 The literature we build on

In the chapters we build on several interconnected strains of the economic literature. Both the first and the fourth chapter builds on the literature on persuasion. In the economic literature there are competing, and sometimes complementary, ways to model persuasion. One approach is a Bayesian one: agents receive messages from different sources, and update their beliefs in light of these messages. Another approach is the one championed by Becker and Murphy (1993) in their paper on advertising: persuasion and the consumption of certain goods or certain activities are best thought of as complements. Persuasive messages can thus shift the demand for certain products or activities, while they may in themselves be regarded as "goods" or "bads".

The literature on persuasion is somewhat connected to the literature on investments in human capital. Just like in the case of human capital investment is general, persuasion often occurs within the family. Parents spend a lot of resources to shape the beliefs, preferences and values of their children. Several papers by Gary Becker have been crucial in paving the way to the study of parental influences on children. Another feature of persuasion is that throughout much of human history persuasion has been

largely mediated by religion. The economic literature on religion, starting with seminal contributions by Iannaccone (see e.g. Iannaccone, 1998), has thus exerted a large influence on our work. Even when studying ordinary product market decisions there are (quasi-)religious motives to consider. Debates around the "right" kind of diet often resemble religious debates. Fortunately, such quasi-religious controversies these days are fought in a peaceful way. A more straightforward application of the economics of religion follows in Chapter 4 where we analyze the use of supernatural claims by parents to influence the behavior of their children, and religious or other supernatural persuasion by a countries' ruler aimed at influencing the behavior of his subjects. Chapter 3 naturally builds on the large body of the literature on human capital. While the distinction between general and firm-specific human capital dominates the research field, the study of industry-specific skills have also gained currency in the past few decades. We build in particular on this strain of the literature.

4 Summary of the thesis chapters

In the first chapter we consider the demand for information in product markets. Traditional analysis held that information is generally positively valued by consumers. Recent work by Loewenstein and others, however, demonstrate "information avoidance" in a number of settings. What is lacking in this "behavioral" literature is the market context: when would misinformation (or self-deception) survive as a market equilibrium and how do certain regulatory and legal changes as well as market incentives effect this equilibrium? In particular, we show that tort liability for (even unwittingly) deceiving consumers increases the demand for new information about the product, while regulation can potentially "save" consumers both from inferior products and the (psychic) cost of finding out that the product is inferior. Nevertheless, we also show that tort law and/or regulation can actually lead to worse outcomes than laissez-faire as they increase the equilibrium price of the product.

The second chapter deals with the effects of the exit of car-sharing company Uber from the Hungarian market. Some of the effects are fairly obvious (such as an increase in demand for ordinary taxi services), but much less has been written about Uber's effect on alternative modes of transportation. In the chapter we exploit the exit of Uber as a natural experiment to study its effect on bicycle sharing usage. Our main findings are somewhat counterintuitive: we find that among pass (season ticket) holders, bicycle usage decreases with the exit of Uber, while among single ticket users it increases. Our conclusion is that there is a kind of "temporal complementarity" between Uber and cycling, in that city dwellers might be more willing to use the bicycle sharing system instead of e.g. their

own cars during the day if they can expect that they can order an Uber ride late in the evening. The exit of Uber thus seemed to have led to a decrease in cycling, and possibly (although we did not study this in this chapter) an increase in own car use.

The third chapter considers human capital investments by parents in their children. When categorizing different human capital investments, traditionally it is the difference between general and firm-specific human capital that has been emphasized in the literature. Yet, as is shown by e.g. Derek Neal (1995), distinguishing industry (or occupation) specific human capital from truly general and firm-specific human capital is also fruitful. The chapter builds on such work in exploring the following, so far underresearched issue: how early and how much do parents optimally invest in occupation specific human capital for their children? The rationale for earlier investments is that there is dynamic complementarity (see e.g. Cunha and Heckman, 2007) in the production of human capital: skills build on earlier accumulated skills. Yet, if there is substantial uncertainty over which occupation would best "fit" the child, and in particular, where the adult child might earn "rents", the parent may delay specialization, investing instead in general human capital. In the chapter we derive several testable predictions. One is that both child prodigies *and* relatively "bad" students specialize more early, while general "good" students typically delay specialization. This is due to the fact that good students are more uncertain in which sector they may earn a rent, while relatively poor students typically cannot earn a rent anywhere (they do not have scarce talents) so they might as well choose a field of specialization randomly. Another implication of our model is that a general increase in market size incentivizes early specialization in areas such as acting, singing, sports or chess, however, by increasing potential rents in all industries, it also leads to a higher level of general human capital investment. Conversely, the significance of early specialization in the form of early vocational education decreases in market size. Another important implication of our model is that socially "manufactured" rents and barriers to entry in professions could result in higher total income than an "intermediate" system of privileges. An "imperfect" system of restricted mobility across industries can be inferior to either fully equal opportunity and a perfect system of privileges. Full equal opportunity, however, is first-best efficient. Finally, the essay is also a contribution to the literature on the sexual division of labor and associated "gender roles". I show that "socially constructed" gender roles or expectations can be efficiency-enhancing as they make it more certain that a particular individual will work in the household (market) sector, leading her or him invest more in household (market) specific human capital, and, if the two are complements, also in general human capital. At the same time, strict societal roles lead to talent misallocation. The higher the costs from talent misallocation and the lower the benefits from a sexual division of labor, the more likely it is that societies will move away from

strict gender roles toward a more "equal opportunity" arrangement.

The fourth chapter considers forms of persuasion that rely on supernatural, such as religious, claims. One example we consider is parents fostering a belief in Santa Claus in their children. In our model, parents do this in order to induce their children to "behave well". The figure of Santa Claus plays the role of a commitment device: parents themselves cannot easily commit to withdrawing gifts and other benefits from their children if they behave badly. Furthermore, they can only imperfectly monitor their child's behavior. A belief in Santa Claus helps to solve this problem: Santa Claus is an impartial figure, who does not love the child the same way the parent loves, so he can easily commit to a reward-or-punish strategy. Santa Claus is also said to be able to monitor the child's behavior in a costless or low-cost manner. Our model sheds light to the fact that many children lose their belief in Santa Claus after they start school. This may be due to interactions with older, better informed peers. Children will likely believe such peers when they deny the existence of Santa Claus as, unlike the parents, they cannot gain by lying about it. The incentive to foster a belief in Santa Claus increases in income as richer parents optimally buy more expensive gifts to their children. In the rest of the chapter we extend our framework to consider supernatural persuasion in the political sphere. We posit a leader who can increase the tax base (induce people to work more instead of engaging in theft or other forms of rent-seeking) in two ways: they can spend on "traditional" law enforcement and they can invest in persuading citizens, using religious doctrine, that theft and rent-seeking in general is "sinful". Persuasion can work even when the probability of apprehension is low as the threat of punishment is "inside the citizen's mind". Religious persuasion commonly occurs through "the Church", and the political leader has to offer part of the tax revenue to the Church in exchange of the Church producing persuasion that increases the tax base. The Church's position is similar to that of a modern television network: it provides services that are valuable from the churchgoers' point of view, but they bundle these services with messages incentivized by third parties.

In the last chapter we propose a few "small" models that deal with the formation of values and preferences in the family. In particular, we study how parents invest in certain character traits or "virtues" such as thrift, self-control or honesty. Such character traits become more important as the cost of monitoring individuals' behavior in certain contexts becomes more expensive. The chapter examines character formation in relation of dealing with addiction and other bad habits, engaging in crime, as well as principle-agent type situations where honesty can serve as a commitment device. We examine two broad cases: in one case, parents are paternalistic toward their children and they invest in instilling preferences for "merit goods". We show however, that parents may still e.g. push their children away from drugs and other harmful habits even if they are fully altruistic, as

first, they can save on giving transfers to their "broke" adult children, and second, when they expect their adult child to shun bad habits, they face a higher return on investing in their child's human capital.

5 Acknowledgments

This collection of essays is a result of years of reading, thinking, writing and fruitful conversations. Here I would like t thank those who helped me the most, be in the form of advice, guidance, comments, questions, conversations or just plain encouragement. Most of all, I am grateful to my thesis supervisor, Barna Bakó for his guidance and for being the coauthor of some of these chapters. I thank Kevin Murphy for helpful comments, guidance and conversations and Glen Weyl for fruitful conversations and correspondence. I thank colleagues at the Corvinus University of Budapest for conversations and comments. I am grateful to Tyler Cowen for always providing quick and on-the-point comments on my drafts. I thank Miklós Rosta and Róbert Somogyi for comments, criticisms and suggestions on the draft version. I thank Dia Horváth for technical help on our proofs and figures. Finally, my late mother was a constant source of encouragement during my PhD studies, for which I am forever grateful. Human capital builds on already accumulated human capital, and my mother provided me with a very valuable stock.

Chapter 1

Heads in the Sand: Information Aversion in a Market Context

1 Introduction[1]

In this chapter we examine the issue of misinformation in product markets from a perspective different to those taken so far in the economic literature. Most of the literature has been exclusively concerned with the incentives of producers to misinform consumers, while consumers themselves have typically been assumed to be exogenously deceivable, or irrational. In this essay, we show that misinformation can be an equilibrium even in the case of consistently utility-maximizing consumers. This is because consumers might form some attachment for the product they consume which makes it costly for them to learn damaging information about it later on. In other words, consumers have incentives to avoid information about certain products.

Even though self-deception and information aversion is a feature of human nature described since ancient times[2], rational information aversion in product markets is an underdeveloped area of economics. Our essay purports to fill part of that void in the literature. In our model, consumers sometimes regard information as a bad, not as a good. This is because individuals, ceteris paribus, dislike being proven wrong in their prior beliefs. If such a taste against more information exists in certain settings, it implies that it can be rational to remain ignorant in the course of a number of consumption decisions. Information avoidance, or the willingness to pay to avoid information stems not merely from the fact that the new information might be unpleasant, but from the fact that

[1]We thank Tyler Cowen, Gergely Csorba, Attila Tasnádi, participants of the Jornadas de Economia Industrial 2018 in Barcelona as well as several participants at the Conference of the Doctoral School of Economics at Corvinus University of Budapest for useful comments, criticisms and suggestions.

[2]Plato has referred to it, as well as other antic sources. It has also featured prominently in the writings of some classical economists, including Hume and Smith.

the new information might be detrimental to the consumer's self-image. Consumers for instance might be confident in their judgment of product quality, so when it turns out that their judgment was wrong, they experience a negative effect on their self-image. Hence they try to avoid such information unless the benefits of the new information outweighs the expected self-image cost. Take the case of healthcare consumption. Information avoidance is a well documented behavior in this field (see e.g. Dawson et al., 2006; Grusky et al., 2007; Sweeny et al., 2010; Howell and Shepperd, 2012a; or Howell and Shepperd, 2012b). In general, individuals may not welcome information about the efficacy of their chosen treatment, medicine or doctors and might opt to remain ignorant about them. Another good example might be the attitude of professors to remain ignorant towards the evaluation of their courses. Even though it might be useful to get feedback about their teaching practices, the acquired new information may also hurt their self-image. In order to prevent the potential psychological cost of unwanted information, it may be rational to delay or prevent the acquisition of new information. We believe that our assumptions regarding consumer preferences over information have some advantages over assumptions usually held by those in the behavioral literature, such as bounded rationality. Under bounded rationality, producers would still have incentives to provide information to consumers in ways that consumers can understand it. However, if consumers refuse to listen to even (in monetary as well as in cognitive terms) free information, that will not be the case.

The, so far limited, economic literature on information aversion provides empirical evidence of its abundance. Ganguly et al. (2017) provide experimental evidence that individuals are willing to pay to avoid being tested for serious illnesses. In politics, as Cowen (2005) writes, such information avoidance is especially abundant. For instance, people who do not believe in global warming might avoid news on that topic, while environmentalists who believe in the most extreme scenarios avoid news sources that are more optimistic about the effects of climate change. Information as a bad also appears in the economic literature of cognitive dissonance, most of all, in the seminal paper by Akerlof and Dickens (1982) who show that when workers can self-deceive about working conditions, a higher level of labor regulation can be optimal than it is generally assumed in the literature. Yet, to our surprise, not much have been written on the role of information aversion in ordinary product markets. In this chapter we examine both the demand and supply side of a distaste for more information by postulating that consumers receive disutility when they are proven wrong in some of their beliefs, while producers, knowing this, can get away with marketing products that are the results of unsuccessful quality improvements.

Our essay is also related to some broader strains of the economic literature. It is

related to the literature of misinformation in general, and through this, to the industrial organization literature on advertising, starting with the seminal articles of Dixit and Norman (1978) and Becker and Murphy (1993). The economics of identity and religion are another two areas related to our approach in various ways. First, investing in *identity capital* (Akerlof and Kranton, 2010; Becker, 1996) might increase the incentive of consumers to self-deceive themselves over the quality of some product which they choose to identify with. Second, such identification has some resemblance to religious beliefs. Religious markets have been modeled in economics as akin to ordinary product markets (see e.g. Iannaccone, 1998), while our research strengthens this association, but with a spin, emphasizing the religious aspects of ordinary product markets.

Another literature close to our approach is the economics of the media and persuasion through the media. Gentzkow and Shapiro (2006) build a model of media persuasion in which consumers have rational preferences but misinformation occurs in equilibrium. By contrast, Mullainathan and Shleifer (2005) consider a model in which news consumers have a preference for receiving news that validate their priors. Their approach is quite close to those taken by us in this essay.

The essay contributes to the economic literature on how to deal with externalities and asymmetric information. The Chicago school of Law and Economics, associated with Posner (1973) and Stigler (1966), have traditionally held that when private ordering is not feasible (for isntance due to high transaction costs), societies should rely primarily on tort law and not on regulation. This is because tort law can directly internalize externalities, while regulation is a blunter instrument: it requires state control of an activity. Regulation has also been shown by Stigler (1971) to often favor entrenched interest groups who can capture regulation and use it to their own advantage. This became to be known as the interest-group theory of regulation, in contrast to the public interest view of regulation, which held that the increase in reliance on regulation in the late 19th and early 20th century has been an efficient response, driven by public demand, to new problems that arose to increasing market concentration and scale economies. Recently, the public interest view has enjoyed something of a comeback in the economic literature. Glaeser et al. (2001) or Glaeser and Shleifer (2003) argue that when corruption is prevalent in the justice system, torts, often set at high rates in order to deter wrongdoings, can be ineffective, while small fines by regulatory authorities can be welfare-enhancing. This chapter adds another modification to the standard view, but it does so not by relying on the assumption that consumers are irrational in some ways, but extending the rationality framework to include preferences toward information. In our framework, simply providing information to consumers, or providing means of protecting their rights through tort law may be ineffective and not first-best efficient even in cases where physical, as opposed to psychic

information costs are negligible. However, our essay also adds one more justification for using tort law as opposed to regulation: the possibility of suing firms for misinformation provides an incentive to consumers to heed to new information about the product, leading to more informed consumers and thus more innovation by firms in the first place. Still, even under tort law not all consumers may choose to be informed, leaving room for regulation to be the most efficient tool in certain cases. More surprisingly, laissez-faire is sometimes optimal from a total welfare perspective even if the court system operates effectively.

The remaining of the chapter is organized as follows. Section 2 presents the model and its main results. In Sections 3 we consider the possible policy responses and their implications. We outline and discuss the possible applications in Section 4 and conclude in Section 5.

2 The Model

Consider the following three period model. In the first period a monopoly firm chooses its level of innovation, which will affect the probability of developing a successful quality innovation. In the second stage consumers choose whether or not to consume the product, given their reservation price and the price of the product. and learn information about the product they have consumed. After this, consumers decide whether they wish to listen to the information and update their prior belief about product quality accordingly. Then, in the last stage they decide whether to continue to consume the product they consumed in the previous period or exit the market.

We assume a mass of consumers normalized to 1, each with a unit demand and deriving a utility of $v(A)$ from consumption, where $A \in \{0, a\}$ denotes two possible values product quality can take with a subjective probability of ϕ consumer attaches to product quality being a instead of 0 and we assume that $v(0) = 0$, while $0 < v(a) \leq 1$. Furthermore, we assume that consumers derive utility (or disutility) directly from the information about product quality. If a consumer receives information that the product quality is lower than her prior assessment, she receives a disutility in form of psychic cost. Let this psychic cost (measured in money-metric utility), denoted by e_i, be a characteristic of the consumers i, and we assume that consumers are uniformly distributed along a unit line with regards of this psychic cost, i.e. $e_i \in [0, 1]$. (We could generalize to $e_i \in [0, 1]$ but it would add complications without sufficiently altering the main results.)

We assume that the utility function is separable in the taste for product characteristics and information about the product. Specifically, we assume that the utility can be written

in the following form

$$U_i(A, e_i) = v(A) - u(e_i) \tag{2.1}$$

where $u(\cdot)$ is continuous and increasing function, while $U'(e_i) < 0$ and $U''(e_i) \geq 0$. In the followings, for the sake of tractability, we assume that the consumer's disutility (more specifically, the monetary value of it) if she learns that product quality is 0 instead of a is e, i.e. $u(e_i) = e_i$.

As we mentioned earlier, there is a monopoly firm that supplies the product.[3] It either produces a high quality product or a low-quality product. The probability that the firm's first period innovation is successful and thus it produces the high-quality product is given by $\psi > 0$. The production costs for any quality realization are, for the sake of simplicity, assumed to be zero. Having a monopoly producer pricing is straightforward. As consumers attach a prior probability of ϕ to the product having quality a instead of 0, the market price in the second period is going to be $p_2 = \phi v(a)$.

When consumers choose to listen to new information, they are delivered either 'good', or 'bad' news about the product. If they receive the former, they increase their valuation from $\phi v(a)$ to $v(a)$. In the latter case they lower their valuation to 0. In this case the consumer leaves the market.

As we have stated earlier, in the second period after consuming the product consumers receive information about the product quality and decide whether to listen to this information. Not listening to new information is a form of information aversion and we can interpret it in various ways. Consumers could literally 'close their ears' to new information, or if they hear new information, they could nevertheless disregard it, clinging to their prior beliefs instead. Based on our assumptions, consumer i.listens to new information if and only if

$$\phi v(a) \geq e_i \tag{2.2}$$

where the left-hand side of the inequality is the benefit of listening to new information, which equals with the price of the product she saves by not consuming the low-quality product. Let e^* denote the consumer, who is indifferent between accepting and not accepting the new information. Thus, we have that $e^* = \phi v(a)$.

We solve the game for its subgame-perfect Nash equilibrium by using backward induction. In the last period, when the quality of the product is already revealed to the consumers who are willing to learn about it, the firm with the higher quality realization either continues to satisfy all of its consumers at a price $\phi v(a)$ or serves only those who

[3]We assume monopoly in order to abstract away from strategic interactions that may arise in case of multiple firms.

are now informed and thus value the product to $v(a)$, at price $v(a)$. The condition for the latter is

$$\phi v(a)^2 > \phi v(a) \tag{2.3}$$

which in never satisfied since $0 < v(a) \leq 1$, thus the firm will always serves all its consumers assuming that produces a high quality product.

Likewise, the firm with the low-quality realization always sets the price of $\phi v(a)$, however, it serves only its captured consumers, those who deceive themselves in the last period. This price yields a profit of $(1 - \phi v(a))\phi v(a)$. Thus, the firm's expected total profit function can be written as

$$E\pi = \phi v(a) + \psi\phi v(a) + (1 - \psi)(1 - \phi v(a))\phi v(a) \tag{2.4}$$

So far we have treated the success of quality improvement as exogenous. In the following we will be tying the probability of successful quality realization to the amount of innovation the firm engages in. For example, a firm might invest in quality improvement technology, however such innovation is not always successful. Suppose the firm can invest in quality improvement at a unit cost $I > 0$. Let the level of investment Z influence the probability of the high quality realization of the product, so that $\psi'(Z) > 0$, while $\psi''(Z) < 0$. Therefore, the firm's expected profit is

$$E\pi = \phi v(a) + \psi(Z)\phi v(a) + (1 - \psi(Z))(1 - \phi v(a))\phi v(a) - IZ \tag{2.5}$$

Maximizing (2.5) with respect to Z, yields

$$\psi'(Z)\phi^2 v(a)^2 = I \tag{2.6}$$

i.e. the marginal benefit of innovation equals the marginal cost.

Naturally the incentive to innovate increases in the marginal product of investment $\psi'(Z)$ and decreases in the investment cost. The incentive to innovate also increases in consumers' marginal valuation of the product quality improvement, i.e. in $v(a)$. Furthermore, it is also increases in ϕ, since a higher ϕ means that more consumers choose to get informed about the product quality and as a consequence the firm earns a lower profit in the case of a bad quality product.

3 Policy responses

In this section we discuss the policy choices that might be available to regulators. Specifically, we consider three possible policy choices: laissez-faire, taxation and regulation.

3.1 Laissez-faire

Total welfare and consumer welfare under laissez-faire can be written as

$$W^{LF} = \phi v(a) + \psi(Z)\Delta_1 - (1-\psi(Z)) \int_0^{\phi v(a)} e \ \mathrm{d}e + \tag{3.1}$$
$$\psi(Z)\left(\phi v(a) + \Delta_2\right) + (1-\psi(Z))(1-\phi v(a))\phi v(a) - IZ$$

and

$$CW^{LF} = \psi(Z)(\Delta_1 + \Delta_2) - (1-\psi(Z)) \int_0^{\phi v(a)} e \ \mathrm{d}e \tag{3.2}$$

where $\Delta_t = v(a) - \phi v(a) = (1-\phi)v(a)$ is the consumer surplus enjoyed by consumers consuming high-quality product in period t, where $t = 1, 2$.

3.2 Tort law

In this section we explore the effects of tort law on consumer and firm behavior. We assume that consumers have the opportunity to sue the firm at a fixed cost s (with $s \geq 0$) when the product is of bad quality. Furthermore, we assume that the court might not decide in favor of the plaintiff even though the product is low quality. The probability that the court rules in favor of the firm in case when it produces a low-quality products is η, where $\eta \geq 0$. The compensation awarded to the plaintiff is equal to the price paid for the product, i.e $\phi v(a)$.

If the consumer chooses not to get informed she will never sue the firm. However, a well-informed consumer might launch a lawsuit against the firm if

$$(1-\eta)\phi v(a) > s \tag{3.3}$$

holds.

Notice that the possibility of starting a lawsuit against the firm alters the incentive of the consumer to learn about the product quality. Thus, with tort law consumer i listens to new information if

$$\phi v(a) + (1-\eta)\phi v(a) - s \geq e_i \tag{3.4}$$

in case when (3.3) holds. On the other hand, if (3.3) is not satisfied, then consumers behave as in the case of laissez-faire and listen to new information as long as (2.2) holds. In what follows we assume that $(1-\eta)\phi v(a) > s$ holds.[4]

From (3.4) follows that the number of informed consumers is $e^* = \phi v(a)(2-\eta) - s$.

The firm's expected profit can be written as follows

$$E\pi = \phi v(a) - (1-\psi(Z))(1-\eta)e^*\phi v(a) + \psi(Z)\max\{e^* v(a), \phi v(a)\} + (1-\psi(Z))(1-e^*)\phi v(a) - IZ \tag{3.5}$$

Simplifying (3.5) yields

$$E\pi = \begin{cases} \phi v(a)[1-(1-\psi(Z))((2-\eta)e^*-1)] + \psi(Z)v(a)e^* - IZ & s < \phi[(2-\eta)v(a)-1] \\ \phi v(a)[1-(1-\psi(Z))((2-\eta)e^*-1)] + \psi(Z)v(a)\phi - IZ & s \geq \phi[(2-\eta)v(a)-1] \end{cases} \tag{3.6}$$

Maximizing (3.6) with respect to Z, we get the following first-order conditions

$$\psi'(Z)v(a)[(\phi((2-\eta)e^*-1) + e^*] = I \tag{3.7}$$

if $s < \phi[(2-\eta)v(a)-1]$,

$$\psi'(Z)\phi v(a)(2-\eta)e^* = I \tag{3.8}$$

if $s \geq \phi[(2-\eta)v(a)-1]$.

Notice, that the marginal benefit from the innovation unsurprisingly decreases in η and s in both cases. Moreover, the optimal level of innovation increases in ϕ and $v(a)$.

Total and consumer welfare under tort law can be given as follows:

$$W^{TL} = \phi v(a) + (1-\psi(Z))\left[(1-e^*)\phi v(a) - \int_0^{e^*} e \, \mathrm{d}e - e^* s\right] - IZ + \psi(Z)\Delta_1 + \begin{cases} \psi(Z)e^* v(a) & s < \phi[(2-\eta)v(a)-1] \\ \psi(Z)(\phi v(a) + \Delta_2) & s \geq \phi[(2-\eta)v(a)-1] \end{cases} \tag{3.9}$$

[4]Notice that we assume that the consumer will not sue the firm if she is indifferent between suing or not suing.

while

$$CW^{TL} = \psi(Z)\Delta_1 + (1-\psi(Z))(1-\eta)e^*\phi v(a) - (1-\psi(Z))\left[\int_0^{e^*} e \ \mathrm{d}e + e^* s\right] + \begin{cases} 0 & s < \phi[(2-\eta)v(a) - 1] \\ \psi(Z)(\Delta_2) & s \geq \phi[(2-\eta)v(a) - 1] \end{cases}$$

where again $\Delta_t = (1-\phi)v(a)$ is the consumer surplus enjoyed by consumers consuming high-quality product in period t, where $t = 1, 2$.

3.3 Regulation

Another possible policy response is quality regulation. The regulator may ban the sale of products advertised as having a quality level a when it judges the actual quality level to be 0. Furthermore, we allow for the possibility that the regulator makes a mistake: it may not ban a product advertised as a although its actual quality level is 0. This can happen for various reasons: the regulator might have imperfect information about the product, or the firm may bribe the regulator to allow its product onto the market.[5] We capture all of these possibilities in a single probability parameter: let the probability of regulatory mistake be λ. Furthermore, we assume that consumers have some level of trust in the regulator and as a consequence update their prior belief that the product is a high quality from ϕ to $\hat{\phi}$, where $\hat{\phi} \geq \phi$. Moreover, we also assume that $\hat{\phi}$ decreases in λ and if $\lambda = 0$, i.e. regulation always screens out the bad product, the consumers will have complete trust in the product, i.e. $\hat{\phi} = 1$. On the other hand, if $\lambda = 1$, i.e. regulation never screens out the bad product, then $\hat{\phi} = \phi$, i.e. the existence of regulation will not affect the consumers' believes about the product quality.

Let us first examine the firm's pricing and investment decisions under regulation. Initially, the firm can always sell any quality at $\hat{\phi}v(a)$. In the second-consumption period the high-quality product is sold on a price of either $v(a)$ (to those who choose to be informed) or $\hat{\phi}v(a)$ (to all consumers). However, notice that $\hat{\phi}v(a)^2$ is never greater than $\hat{\phi}v(a)$, thus as in the case of laissez-faire the firm is always better-off by serving all consumers. Yet, in the second-consumption period if the product is low-quality but the regulator approves it the firm can sell the product only to the consumers who choose to remain ignorant, at a price $\hat{\phi}v(a)$. If the regulator does not approve the low-quality product the firm makes zero profit at this last period.

[5] One source of mistake might be that the product performs differently in the trial period and after it is brought to the market, as it recently turned out to be the case with some car manufacturers.

Thus, the firm's expected profit can be written as

$$E\pi = \psi(Z)(2\hat{\phi}v(a)) + (1-\psi(Z))\lambda[\hat{\phi}v(a) + (1-\hat{\phi}v(a))\hat{\phi}v(a)] - IZ \tag{3.10}$$

Taking the first derivative of equation (3.10) with respect to Z yields the following first-order condition

$$\psi'(Z)\hat{\phi}v(a)[2-\lambda(2-\hat{\phi}v(a))] = I \tag{3.11}$$

Notice that, the higher the probability of mistaken regulation, the lower the incentive of the firm to innovate. Moreover, the effect of $v(a)$ and $\hat{\phi}$ on innovation is positive. Intuitively a higher expected valuation means that more consumers choose to get informed and at the same time if regulation works reasonably well the firm can gain more if it produces a high-quality product. These effects increase the firm's incentive to innovate.

Total welfare and consumer welfare under regulation can be given as

$$\begin{aligned} W^R = \psi(Z)[2\hat{\phi}v(a) + \hat{\Delta}_1 + \hat{\Delta}_2] + \\ (1-\psi(Z))\lambda\left[\hat{\phi}v(a) + (1-\hat{\phi}v(a))\hat{\phi}v(a) - \int_0^{\hat{\phi}v(a)} e \ \mathrm{d}e\right] - IZ \end{aligned} \tag{3.12}$$

and

$$CW^R = \psi(Z)[\hat{\Delta}_1 + \hat{\Delta}_2] - (1-\psi(Z))\lambda\left[\int_0^{\hat{\phi}v(a)} e \ \mathrm{d}e\right] \tag{3.13}$$

where $\hat{\Delta}_t = (1-\hat{\phi})v(a)$ is the consumer surplus enjoyed by consumers consuming high-quality product in period t, where $t = 1, 2$.

3.4 Comparing policy responses

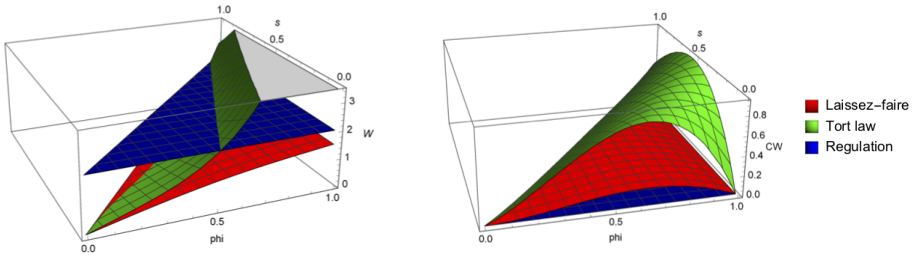
In the following in order to draw policy relevant conclusions we will compare total and consumer welfare under the above discussed policies and derive comparative statics results for some key variables. To make the analysis more tractable we will consider without loss of generality specific functional forms and variable values. In particular, we normalize the value of $v(a)$ to 1 and assume that $\psi(\cdot) = \sqrt{\cdot}$. Furthermore, we assume that $\hat{\phi} = 1-(1-\phi)\lambda$. Finally, we set the unit cost of innovation, i.e. I, to $\frac{1}{2}$.

Let us first consider the case, when both regulation and the court system is completely competent and/or incorruptible. It then follows that $\hat{\phi} = 1$, that is, consumers have perfect confidence in the product (as they, rightly, have perfect confidence in the regulator).

i.e. $\eta = 0$ and $\lambda = 0$.

A well-functioning regulation dominates laissez-faire on total welfare. The main force behind this result is that with an effective regulator only the high-quality product can be sold thus the firm has a very strong incentive to innovate. Furthermore, in this case consumers will never be 'disappointed' by the product. Regulation welfare-dominates tort law under low values of ϕ, while tort law dominates regulation for sufficiently high values of ϕ. This is because, as $\hat{\phi} = 1$, ϕ does not have an effect on W^R but it increases W^{TL} as a higher ϕ implies a larger number of informed consumers. However, for any given ϕ welfare under tort law decreases in s as a higher litigation costs makes tort law less effective in incentivizing innovation, since fewer consumers will choose to get informed. Thus, for a higher ϕ even with a greater s tort law may welfare-dominate regulation.

Regulation leads to lower consumer welfare than either laissez-faire, which follows from the fact that consumers have perfect confidence in the product and the monopolist will take advantage of this in its pricing behavior. Not surprisingly, under tort law consumer welfare is at least as high as with laissez-faire, since consumers will always get compensation if they happen to consume a low-quality product, provided that they choose to get informed and file a suit against the firm. Interestingly, consumer welfare does not always decrease in s. Instead, it first increases, then decreases in it. Even though a higher s means fewer consumers will choose to get informed which eventually leads to worse expected product quality, it also yields a lower psychic cost. Moreover, notice, that consumer welfare monotonically increases in ϕ only where it decreases in s. Otherwise, it first increases then decreases in it. The mechanism responsible for this is similar to the one we have seen with s. These results are illustrated in Figure 3.1.



(a) Total welfare (b) Consumer welfare

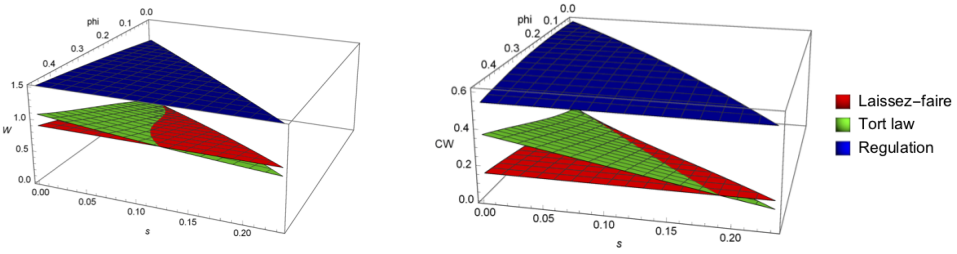
Figure 3.1: Total and consumer welfare if $\eta = 0$ and $\lambda = 0$.

Now let us consider cases where either tort law or regulation works perfectly, however, the other is not effective at all. If, for example, $\lambda = 0$ and $\eta = 1$ then the court system essentially leaves everything unchanged relative to laissez-faire. Hence we essentially compare laissez-faire with regulation. As we had seen it above, regulation is more desirable from a total welfare perspective, yet it is less desirable in terms of consumer welfare. In the

reverse case ($\lambda = 1$ and $\eta = 0$), however, tort law most of the time dominates laissez-faire (in lieu of regulation, which is completely ineffective in this case) on both total welfare and consumer welfare. Interestingly, however, for low values of ϕ coupled with high values of s total welfare is the greatest under laissez-faire. Even though the incentive to innovate is negatively effected by both a low ϕ and a high s, s only takes its effect on tort law. We do not observe this pattern when it comes to consumer welfare. The simple reason for this is that consumers also get compensated for consuming a low-quality product.

Next, consider cases where both regulation and tort law are, to some extent, effective, but highly imperfect. Take, for example, $\lambda = \frac{1}{2}$ and $\eta = \frac{1}{2}$. In this case, consumers will take advantage of tort law if and only if $s < \frac{\phi}{2}$, otherwise to file a case against the firm is prohibitively expensive. Therefore, consumers who choose to be informed under tort law can be given as $e^* = \min\left\{1, \frac{3\phi}{2} - s\right\}$. Furthermore, the innovation level the monopolist is going to implement under tort law is $Z^{TL} = \min\left\{1, \frac{1}{16}(\phi(2 + 9\phi - 6s) - 4s)^2\right\}$, while under regulation it is $Z^R = \min\left\{1, \frac{1}{64}(1 + \phi)^2(5 + \phi)^2\right\}$. As a consequence, we have to differentiate between four subcases:

i) if $\phi < 2\sqrt{3} - 3$. Total welfare is always the greatest under regulation. This is not surprising given the low level of ϕ. Tort law, on the other hand, dominates laissez-faire only for low levels of s. The relative desirability of tort law from a total welfare perspective also increases in ϕ. In essence, if ϕ is relatively low then s has to be low as well in order for tort law to dominate laissez-faire. These second-best options can be important as a decrease in the effectiveness of regulation and/or an increase in its cost may decrease total welfare attainable with regulation. Interestingly regulation is not only the best from the point of view of total welfare but also when it comes to consumer welfare. That is, we arrive at a seemingly surprising conclusion: consumers might gain by regulation only when regulation works imperfectly. See Figure 3.2.

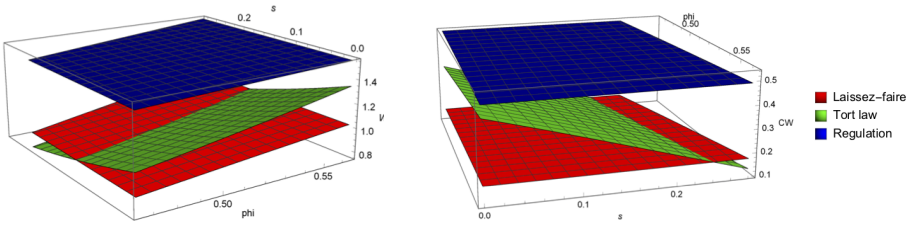


(a) Total welfare (b) Consumer welfare

Figure 3.2: Total and consumer welfare if $\eta = \frac{1}{2}$ and $\lambda = \frac{1}{2}$, while $\phi < 2\sqrt{3} - 3$

ii) if $2\sqrt{3} - 3 < \phi \leq \frac{1}{9}(\sqrt{37} - 1)$, then innovation under regulation at is its maximal level, while under tort law it lower than that. Regulation is always the best from both a

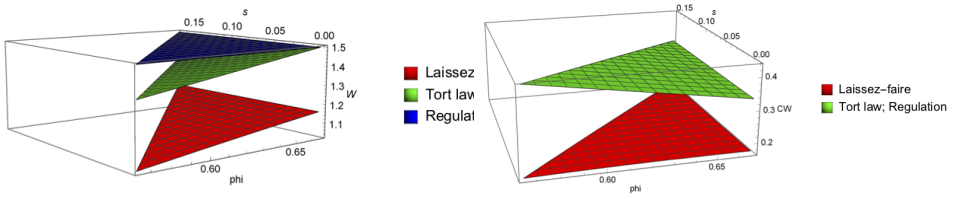
total welfare and a consumer welfare point of view. As s decreases, however, welfare under tort law converges to welfare under regulation, although it never reaches its level. Tort law dominates laissez-faire under sufficiently low values of s. Similar relationships can be uncovered about consumer welfare. The effects of ϕ on welfares are similar as in the previous cases. See Figure 3.3.



(a) Total welfare (b) Consumer welfare

Figure 3.3: Total and consumer welfare if $\eta = \frac{1}{2}$ and $\lambda = \frac{1}{2}$, while $2\sqrt{3}-3 < \phi \leq \frac{1}{9}(\sqrt{37}-1)$

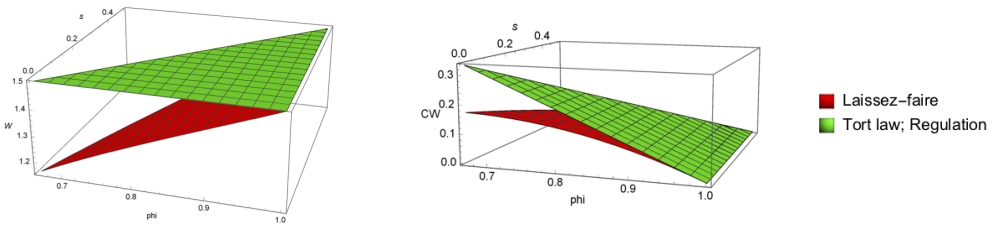
iii) if $\frac{1}{9}(\sqrt{37}-1) \leq \phi < \frac{2}{3}$, then innovation under both regulation and tort law is at the maximum, i.e. $Z^R = Z^{TL} = 1$ even if not every consumer chooses to get informed. Total welfare under regulation is initially strictly greater than welfare under tort law (and laissez-faire), however, as ϕ increases and s decreases, the two converge, and eventually yield the same total welfare. Notice, that this increase in total welfare under tort law is not due to more innovation as that is already at its maximum, rather due to an increase in the number of well-informed consumers. Since the firm serves only the well-informed consumers this results in an increase in the total surplus. When it comes to consumer welfare, regulation and tort law generates the same level. Note, that this holds for every value of s. Given that product quality is high with certainty this might not come as a surprise, since consumers will never exercise the right to sue the firm, therefore they will not incur the cost s. Laissez-faire always generates a lower level of consumer welfare than the other two regimes, due to a lower level of innovation. These results are illustrated on Figure 3.4.



(a) Total welfare (b) Consumer welfare

Figure 3.4: Total and consumer welfare if $\eta = \frac{1}{2}$ and $\lambda = \frac{1}{2}$, while $\frac{1}{9}(\sqrt{37}-1) \leq \phi < \frac{2}{3}$

iv) if $\phi \geq \frac{2}{3}$, then innovation is maximal (which is unsurprising given Z was equal to 1 even at lower values of ϕ, and we have already established that innovation increases in ϕ under each of the possible policies) and every consumer chooses to be informed under tort law. Because of this, welfare under tort law and regulation is the same, i.e. $W^R = W^{TL} = \frac{2}{3}$, which is above W^{LF} throughout, however, welfare under laissez-faire converges to it as ϕ increases. Consumer welfare is even more interesting. Consumer welfare decreases in ϕ under all policies, but decreases faster under regulation and tort law than under laissez-faire. Even though consumer welfare is initially larger under the former two policies, consumer welfare under laissez-faire converges to $CW^R = CW^{TL}$ from below. See Figure 3.5.



(a) Total welfare

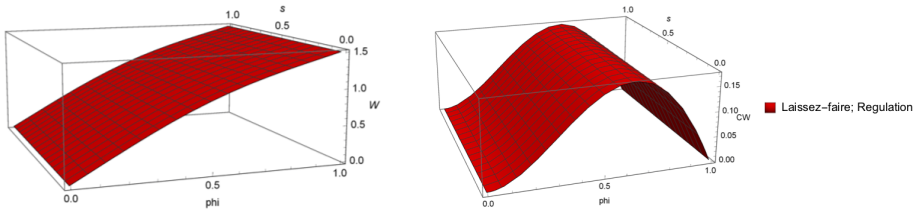
(b) Consumer welfare

Figure 3.5: Total and consumer welfare if $\eta = \frac{1}{2}$ and $\lambda = \frac{1}{2}$, while $\phi \geq \frac{2}{3}$

Finally, let us consider the case with $\eta = \lambda = 1$, that is, both regulation and tort law are completely ineffectual. Under these conditions, no consumer will make use of tort law, therefore we can safely ignore the existence of tort law. Regulation on the other hand generates the same results as laissez-faire, since it allows every type of products into the market. As consumers cannot trust the regulator their initial belief about product quality is unchanged relative to laissez-faire, i.e. $\hat{\phi} = \phi$. Hence, both pricing and innovation is the same under regulation and laissez-faire. However, if we assume some positive fixed cost of introducing and operating a regulatory system, laissez-faire is always the best policy. Total welfare increases monotonically in ϕ, however, ϕ and consumer welfare has a non-monotonic, inverted U-shaped relationship. We have already seen something similar in the first case, and the intuitive explanation is likewise. This is depicted on Figure 3.6.

4 Discussion

The main novelty of our results is that tort law has an additional advantage, beside the advantages discussed in the earlier economic literature: it is especially potent in incentivizing consumers to search for or listen to new information about the product. Hence, while results by Glaeser and Shleifer (2003) on the rise of the regulatory state are



(a) Total welfare (b) Consumer welfare

Figure 3.6: Total and consumer welfare if $\eta = 1$ and $\lambda = 1$.

sound, if the court system is sufficiently well-functioning, tort law can still be better than regulation especially in markets where consumers have incentives to self-deceive. More surprisingly however, we find laissez-faire to be sometimes optimal from a total welfare perspective even if the court system operates effectively.

In our analysis we analyzed a simple model in order to ensure tractability. We can, however, look at some possible modifications and how these modifications might affect our results. We may, for instance, assume a downward-sloping demand curve at the market level or even at the individual level. Suppose, for instance, that each consumer has a downward sloping demand curve, thus, they would like to consume more than one unit of the good. Pricing in such a case is standard monopoly pricing along the market demand curve. At price p the consumer, by listening to new information, saves not only p, but pQ. In this case consumers have a much stronger incentive to be informed. We may, therefore expect more self-deceiving' behavior when consumers, at a given price, consume relatively few units of the good. Another interesting aspect we have not touched on so far is whether the psychic cost of new information is heterogeneous across industries and if so, which industries we should expect to manifest itself the most. Our model considered only one given industry, hence this omission. Still, it might be useful to sketch some intuitions on this issue. The first question that may arise is how consumers' preferences against new information might develop in the case of a particular product. We explicitly modeled consumer behavior so that consumers face the information disutility after they first consumed the product. The reason for this assumption is the intuition that while prior beliefs cannot be observed directly by others, consumption choice can be a useful proxy for beliefs. Hence, if consumers choose differently in the second consumption period, they basically signal others that previously they held false beliefs about the product. Also, we expect the information disutility to be greater when there is a connection between consuming the product and the consumer's personal identity. Consumption of clothes, cars or music, for instance, are heavily tied to identity, and increasingly the same is true for food choices.

We could, as we have indicated in our comparison of individual policies, incorporate

into our analysis the cost of introducing and operating a regulatory system that deals with misinformation. If such costs are substantial, then in some of the cases where we found regulation to be optimal it might not be optimal after the inclusion of these costs.

A further generalization of the model would be to consider an oligopoly instead of a monopoly in order to capture the strategic interactions present in such markets. Allowing for strategic interaction, however, would significantly complicate the analysis.

Another extension would be introducing more than two consumption periods, instead of just two. We could further enrich this model by assuming that over time consumers accumulate consumption capital (as in Becker and Murphy, 1988) that influences the disutility they receive from receiving bad news about the product. Intuitively, the more a consumer has already consumed from a product, the greater pain she will feel if it turns out the product is of low quality.

More formally, suppose that the disutility from new (negative) information depends on a consumption stock or consumption capital S, which is built up in previous consumption periods, so that we have a function $e(S)$, where $e'(S) > 0$, and $e''(S) \leq 0$. S may also influence the marginal utility of consuming the product, as in rational addiction models, but here we are more interested in its effect on the disutility from information about the product. In building this extended model we may consider consumers to be myopic with regard to how their present consumption effects their future information costs, however, in our view it is both more realistic and more instructive to assume that they are aware that they are accumulating consumption capital and that will have an effect on their decision of getting or not getting informed. If consumers know that as they increase their consumption of a product, they will become less likely to acquire information about it later on, they will optimally be less trusting toward the product, that is, they will have a lower ϕ. Furthermore, they will have an incentive to procure information early on. Note also that at period t the consumer, if she listens to new information and the product turns out to be low quality, saves $\beta^{n-t}\phi v(a)$, where β is the discount factor. An implication of this is that older consumers and those with terminal illnesses might be less willing to listen to new information, as they rationally expect to consume in a smaller number of remaining periods. Older consumers presumably also have more consumption capital, which again makes them less likely to listen to new information about the product.

Although our essay focuses on ordinary product markets, a particular market where we would expect rational self-deception to be an especially potent player is the political market. Individuals there have a very low marginal effect as voters on political outcomes, so they can stick to their prior beliefs at a very low price.

Finally, another extension could be to consider naive-Bayesian consumers. In such a case if, for instance, regulation is completely ineffective, consumers might still believe,

with some probability, that it is competent, therefore $\hat{\phi}$ would not necessarily be equal to ϕ. It would likely modify our result to some extent, but not our general qualitative predictions.

The results and extensions presented in this article apply regardless of whether the information costs arise due to physical or psychic reasons, that is, regardless of whether consumers face a standard information cost or a disutility when obtaining information. We stressed the role of disutility because in most cases, even if consumers face information costs, there would be businesses or other parties who would have an incentive to provide free information to consumers. In many settings information problems persist only if consumers are unwilling to accept or act on new information. This is, in short, why we focused on this (admittedly special) case in this essay.

5 Conclusion

In this essay we have analyzed the effect of information avoidance in a market setting. We show that consumers' tendency to stay uninformed can persist even with zero physical information costs. However, market as well as institutional forces can dampen the effect of information avoidance. In particular, stict tort liability or ex ante product regulation can increase welfare, however, relying only on market forces (i.e. laissez-faire) can lead to better outcomes in certain cases. Our findings add to the growing comparative literature on tort law and regulation, as we argue that one advantage of tort law, not emphasized in the literature so far, is that it increases consumers' incentives to get informed.

Chapter 2

Does Uber Affect Bicycle-Sharing Usage? Evidence from a Natural Experiment in Budapest

1 Introduction [1]

Bicycle-sharing systems (BSS) are gaining popularity in more and more cities throughout the world. There is substantial research concerning the utilization of these systems to better understand their success and failure factors. In this study we contribute to this discussion by analyzing the interplay among BSS and other transportation options in the case of Budapest. Recently, a number of changes have taken place in the transportation system of Budapest, e.g. launching a bicycle-sharing system, or the market entry and exit of Uber. These changes give us the possibility to analyze and understand the role Uber may play in local transportation. Specifically, we exploit a legal change that occurred in July 2016 in Hungary that caused the exit of Uber from Budapest to analyze the effect of Uber on BSS usage. From the point of view of our research the relevant question is whether Uber and bike-sharing are substitutes or complements. In case they are substitutes one would expect the exit of Uber to increase the demand for BSS, while if they are complements one would expect the opposite. Our findings suggest a complementary relationship. We find that Uber leaving the market decreased BSS usage

[1]This chapter was published as 'Does Uber affect bicycle-sharing usage? Evidence from a natural experiment in Budapest", *Transportation Research Part A-Policy and Practice*, 133 pp. 290-302., 2020, with Barna Bakó, Zombor Berezvai and Enikő Vigh. The research was supported in the Széchenyi 2020 program framework (EFOP-3.6.1-16-2016-00013) under the European Union project titled: "Institutional developments for intelligent specialization at the Székesfehérvár Campus of Corvinus University of Budapest". The authors thank the BKK Centre for Budapest Transport for access to the data and Péter Dalos for his helpful comments and suggestions. The first author gratefully acknowledges financial support from the National Research, Development and Innovation Office (FK-132343).

overall. This effect comes from regular users of BSS who use the service with passes. For individuals using the service with single tickets the effect of Uber's exit has a positive effect, suggesting substitution across the two services. The results indicate that consumers who are more likely to buy bicycle-sharing passes use bicycle-sharing and Uber as parts of a multimodal way of transportation. Multimodal transportation is a transportation usage pattern characterized by users using a combination of different modes of transportation in a sequential manner (see e.g. Crainic and Kim, 2007). Thus, a person may use public transport, Uber and BSS in a complementary manner to reach her destination. However, if Uber is not available they may use their own car insted.

Our reasoning is supported by a number of previous findings. Several studies have found that BSS is popular mainly among young, urban, college educated and higher income individuals, who are not captive to public transport (see Raux et al., 2017, Goodman and Cheshire, 2014, Ricci, 2015). Raux et al. (2017) further finds that the majority of the users hold a driving license or otherwise have access to a private car. Similar demographic attributes have been found for Uber users as well (see e.g. Hall et al., 2017). Furthermore, these individuals in general tend to use the transportation system in a multimodal way as shown by Clauss and Döppe (2016) or Olafsson et al. (2016).

Based on these findings it is likely that a large number of individuals are able and willing to use both services. Moreover, Hall et al. (2017) shows that Uber can become an integral part of a multimodal transportation system.Taking away Uber as an option therefore might decrease the demand for other travel modes, including bike-sharing, since individuals might prefer to use their own car instead of choosing public transport as a substitute when Uber is no longer available. This line of reasoning is supported by Hampshire et al. (2017) who find that after the suspension of Uber and Lyft services in Austin, TX, 45% of the population surveyed switched to the use of personal vehicles while only 3% shifted to public transport.

Our findings suggest that the nature of complementarity between ride-sharing and bike-sharing services is best characterized as a type of temporal complementarity. Many consumers use these two services at different times of the day, limiting substitution across them. To put it in a more intuitive manner, consider an individual who travels to the city center and wishes to stay there until late at night, perhaps to engage in "partying". If Uber is available, this city dweller can leave her car at home. She can use bike-sharing to reach her intended destination during the day and can use Uber late at night to get back home. However, if Uber is not an available service, she might use her own car to get in and, perhaps by hiring a driver, to get out of the city center. At least, this seems to be the case with many pass holders. For individual ticket holders, Uber and bicycle-sharing appear to be substitutes. Consider, for example, a tourist who ponders whether to buy

a BSS ticket but then, perhaps out of convenience, orders Uber instead. A substitutory relationship in such cases seems intuitive.

Our findings of net complementarity receive further support from the fact that Uber has recently set up a bike-sharing service, JUMP (see www.uber.com). If ride-sharing and bike-sharing are substitutes, it would not be a profit-maximizing strategy from Uber's part, however if they are complements, it could be. Establishing complementary services increases demand for both services.

Uber has made it into the headlines many times since its launch. Besides its novel concept of transportation, most often the news has been about the controversy of "not playing by the rules". We are not aiming to take sides in this debate, however our results suggest that the presence of Uber (and Uber-like services) could have significant spillover effects on passengers' behavior related to other means of transportation. This may put the way we think about Uber in a different perspective that might be worth taking into consideration in future policy decisions.

The structure of the article is as follows. In Section 2 we provide some background on bicycle-sharing, Uber and similar ride-sharing services available in Budapest, as well as the Hungarian regulatory changes we have studied. In Section 3 we introduce our dataset and we present our empirical model. We discuss our results and draw policy implications in Section 4. The chapter is concluded with a summary in Section 5.

2 Background

2.1 BSS in Budapest

The bicycle-sharing system of Budapest (called MOL Bubi) was established in September 2014. In the first round, 76 stations were opened. The system was expanded in three phases and reached 112 stations by the end of 2016 (see Figure 2.1 for the geographical location of the stations). Currently, the system densly covers the inner area of the city.

Users can buy quarterly, semi-annual and annual passes or 24-hour, 72-hour and weekly tickets. Both the tickets and the passes allow unlimited number of bicycle hiring. The first 30 minutes of each rent is free of charge. Renting a bicycle for longer than 30 minutes comes with an additional variable fee that depends on the length of the usage.

The 24-hour ticket costs around €1.6, while the price of the annual pass is around €60. The additional variable fee is €1.6 per 30 minutes. Passes have two favorable features compared to tickets. First, one pass allows the use of four bicycles at the same time, i.e. groups of people can buy only one pass together and share the related costs. Second, 15-25% of the price of the passes can be used to cover the variable fee (that applies for a

rental lasting longer than 30 minutes). Prices have not changed since the launch of the system, therefore they could not influence the change of the usage patterns.

The daily operation of the BSS is managed by a third-party company. This company is penalized based on the number of empty and full stations. To be more precise, the penalty applies if there are less than two bicycles or less than two empty places available in any station, which, however, rarely happened according to the data.

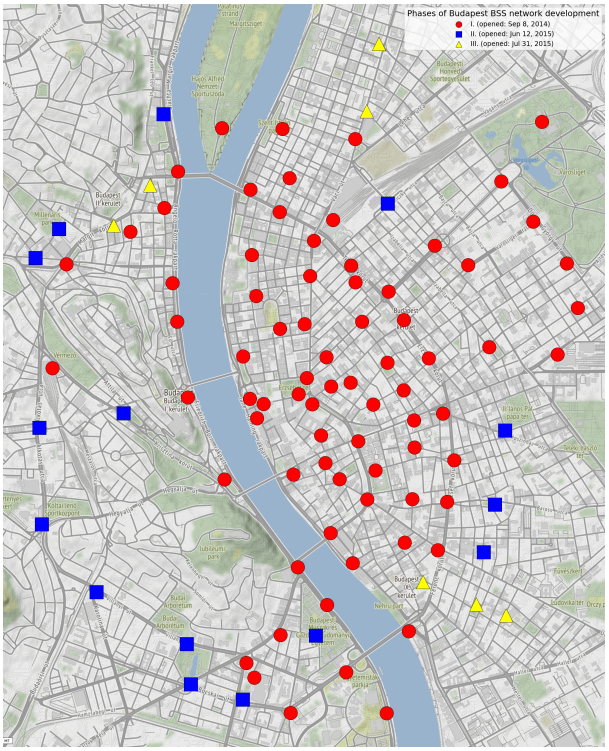


Figure 2.1: BSS station development in Budapest (size of the dots represents number of bicycle docks at a given station)

2.2 Uber

Uber is a highly valued start-up company that provides taxi-like transportation services currently in more than 700 cities worldwide. Uber prices are significantly lower compared to taxi prices, however, prices depend on supply and demand and this flexibility could cause very high prices in some peak periods (e.g., at New Years' Eve). The service can be ordered via a mobile application available for the three most widespread smartphone operating systems (iOS, Android and Windows Phone). After downloading and registering for the application, potential passengers can check the location of the closest available Uber cars. They can see the rating and reviews of the driver and decide which car they would like to order. The driver can also check the profile of the potential passenger and

confirm the order. After the trip, payments are done via the mobile application, i.e. there is no cash transaction in the cars. Finally, both the passenger and the driver can evaluate each other to generate additional ratings and reviews. The system of Uber is considered an innovative one and this is a major factor behind its success in several cities.

Uber drivers are self-employed contractors using their own cars and driving whenever they want. Since this activity is mostly outside of the scope of taxi regulations, several protests have taken place against Uber's operation around the world. Protesting taxi drivers indicate unfair competition between unregulated Uber and regulated and licensed taxi services. This has prompted some cities to regulate more strictly the operation of Uber.

Uber was launched in Budapest in November 12, 2014, and, as was the case in most cities, drew much controversy. The main critics of the ride-sharing service came from local taxi driver associations that accused Uber of "not playing by the rules". There was truth in that accusation. In 2013 Budapest introduced a new taxi regulation. The most important part of it was the introduction of a mandated price. Uber provided similar services for, on average, less than half of that price. Lower prices caused fast penetration for Uber and triggered intense protest from taxi drivers against the company. The campaign, which featured demonstrations by taxi drivers, was in the end successful. The Hungarian parliament accepted a new law that made Uber's operation in Hungary (in essence Budapest) impossible. Uber quit the market on the day (July 24, 2016) the law came into force.

3 Data and Methodology

BSS related data were provided by the system operator, Centre for Budapest Transport. The dataset contains start date, end date, start station, end station, and ticket type (pass or ticket) for all the trips occurred in 2015 and 2016. Usage patterns show significant seasonality (see Figure 3.1), BSS is much more utilized during summertime. Since the exit of Uber happened in the middle of summer (July 24, 2016), we decided to use the summer periods only, i.e., from June 1 to August 31 for both years. This shorter sample makes it possible to analyze the most utilized periods. Additionally, the shorter period enables a regression discontinuity-type of analysis that is often used in treatment effect identifications (O'Keeffe and Baio, 2016) to mitigate the unobservable changes that might occur in a larger time window.

The dataset allowed us to separate users based on ticket types, that is, to differentiate regular users (who are using the service with passes) from ad hoc users (who are using the service with tickets). Some data cleaning was required to eliminate invalid entries. If

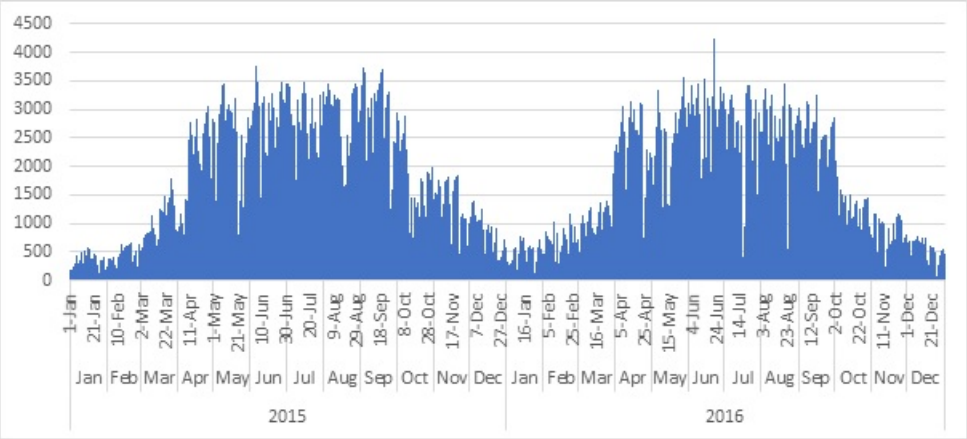


Figure 3.1: Daily usage frequencies of the Budapest BSS (total number of trips per day)

a trip was no longer than 1 minute or either the start or the final station was missing, the trip was deleted from the database. After this cleaning, 511,539 trips remained in our database. The majority (85%) of the usage was generated by regular users and only 15% was connected to tickets. Furthermore, the service is more frequently used on weekdays, and only 25% of the total usage is connected to weekends (Saturdays and Sundays) (see Table 2.1). This is in line with previous findings (see e.g. El-Assi et al., 2017; Faghih-Imani and Eluru, 2015), that indicate that weekdays and weekends might show different dynamics and usage patterns. Usage on weekdays is more connected to commuting, while weekend usage is more about leisure and recreation. The dataset was, therefore, separated into weekdays and weekends samples. The utilization of the stations varies heavily. While the average trip generation was 29.3 per station per day, it is ranging from 0 to 148. Trip data were summarized into number of trips by day, generating station and ticket type.

Ticket Type	Weekday	Weekend	Total
Pass	336,400	98,334	434,734
Ticket	49,771	27,034	76,805
Total	386,171	125,368	511,539

Table 2.1: Number of trips for the summers of 2015 and 2016

Table 2.2 reports summary statistics of the data used. It shows that regular users use BSS more often on weekdays, which can be attributed to commuting to work. On the other hand, ad hoc users use the service more frequently during weekends.

To identify the causal effect of Uber on BSS usage, a control and a treated group need to be compared. However, there is no natural control group available, and therefore, it was required to construct a counterfactual. In this study, we exploit the fact that Uber

Variable	Obs.	Mean	Median	Standard deviation	Min	Max
Number of trips per station with pass on weekdays	12,496	26.9	23	16.9	0	144
Number of trips per station with pass on weekends	4,950	19.9	16	16.2	0	148
Number of trips per station with ticket on weekdays	12,496	4.0	2	5.8	0	61
Number of trips per station with ticket on weekends	4,950	5.5	2	7.6	0	69
Number of stations	184	95.2	98	5.4	76	99
PET scores (hourly data)	4,416	18.4	17.5	6.7	5.7	36
Total daily precipitation (mm)	184	2.5	0	7.8	0	66

Table 2.2: Summary statistics

was available in the whole summer of 2015, but its service was terminated in the middle of 2016. We use the data of 2015 as a counterfactual for 2016. The first half of the summer of 2016 enables us to identify the usage differences between the two summers, and thus, estimate the impact of Uber as a treatment effect. We created the difference between the 2015 and 2016 data to analyze the changes between the two summers. More specifically, since subtracting the same day (e.g., July 1, 2015 from July 1, 2016) might cause a bias in comparing a weekday to a weekend day, we always subtracted the same types of days from each other (i.e., a Sunday was subtracted from the closest Sunday a year before). In this way we capture the changes in trip generation by station, day and ticket type between the two summers.

To account for the differences between the two consecutive years, we control for the most important variables affecting BSS usage based on prior literature. There is a consensus (see e.g. Saneinejad et al., 2012; Gebhart and Noland, 2014; El-Assi et al., 2017; De Chardon et al., 2017) that weather conditions (e.g. temperature, wind speed and precipitation) have major effects on BSS usage. We use Physiological Equivalent Temperature (PET) scores in order to capture the effect of thermal related weather conditions on BSS usage. PET is one of the most commonly used thermal indicator for assessment of the thermal conditions (mean radiant temperature, air temperature, humidity and wind speed) of the human body (see Matzarakis et al., 2007 and Matzarakis et al., 2010). We used RayMan 1.2, developed by the Meteorological Institute, University of Freiburg, Germany, which is a micro-scale model to calculate radiation changes in different environments, to calculate the PET scores. We set geographic longitude at 19°2' and latitude at 47°30', altitude at 105 m and time zone at UTC+2.0 representing the

geographic parameters of Budapest. Furthermore, the average weight was set to 83 kg and height to 176 cm (the average weight and height of Hungarian males[2]; HCSO, 2018). Physiological parameters were constant with an internal heat production of 80 W and a heat transfer resistance of the clothing of 0.9 clo. Since precipitation is not considered in the calculation of PET scores, we control for it separately. The effects of thermal conditions and precipitation on bicycle usage are not linear, therefore several intervals were created from PET scores and precipitation data. Weather data were obtained from the European Climate Assessment & Dataset provided by the European Climate Support Network.

Naturally, ticket and pass prices also affect usage (see e.g. Goodman and Cheshire, 2014; Fishman, 2016; Lin et al., 2017). Yet, these prices have not changed since the launch of the service. Moreover, taxi prices and public transportation prices did not change either in the analyzed period, thus, we did not include price-related data in our analysis.

Another important factor that might influence usage is the size of the BSS network (see e.g., Gebhart and Noland, 2014; Campbell and Brakewood, 2017). Therefore we also controlled for this factor in our regression. Furthermore, several studies suggest that changes in natural and built environment, in public transportation routes, temporary traffic constraints, etc. might affect the utilization of a BSS station (see e.g., Nair et al., 2013; Fishman et al., 2015; Mateo-Babiano et al., 2016; Wang et al., 2016; Noland et al., 2016; Gonzalez et al., 2016; Faghih-Imani et al., 2017b). To account for these changes, we included station-specific fixed-effects in the regression. Finally, to account for the between-day variations, we included day of week dummy variables in the model.

Our model can be written in the following general form:

$$\Delta y_{it} = \beta \Delta Uber_t + \Gamma \Delta x_{it} + c_i + u_{it} \tag{3.1}$$

where y_{it} is the total number of trips generated by station i on day t; $Uber_t$ is a dummy variable taking the value of 1 if Uber was available in Budapest on day t and 0 otherwise; x_{it} contains all the control variables and c_i captures station-specific effects. The unexplained random error term is represented by u_{it}. The model was estimated using fixed effect panel regression. We assumed AR(1) error term in the fixed effect regressions in the weekday subsamples. This is because our dataset contains daily observations that can cause autocorrelation in the dependent variable. However, for weekends, normal fixed effect model was used, as autocorrelation is not relevant in that case, since the weekend subsample contains only two consecutive days per weekend.

[2]The predominant user group of the BSS.

4 Results and discussions

The analysis is divided into three parts. First, we look at the average temporal trends of the network (Section 4.1); second, we analyze the differences of the two summers using panel regression methods (Section 4.2); finally, we divide the sample into five time periods to capture the daily temporal differences using a panel regression framework (Section 4.3).

4.1 Usage patterns

We begin our analysis by considering the unconditional changes that happened after the exit of Uber. Figure 4.1 and 4.2 depict the average trip generation per station for different ticket types during weekdays and weekends when Uber was present in Budapest and after its exit. The figures show data for the summer of 2016 (June 1-August 31). Since no new station was added to the network in this period, network expansion does not bias the data.

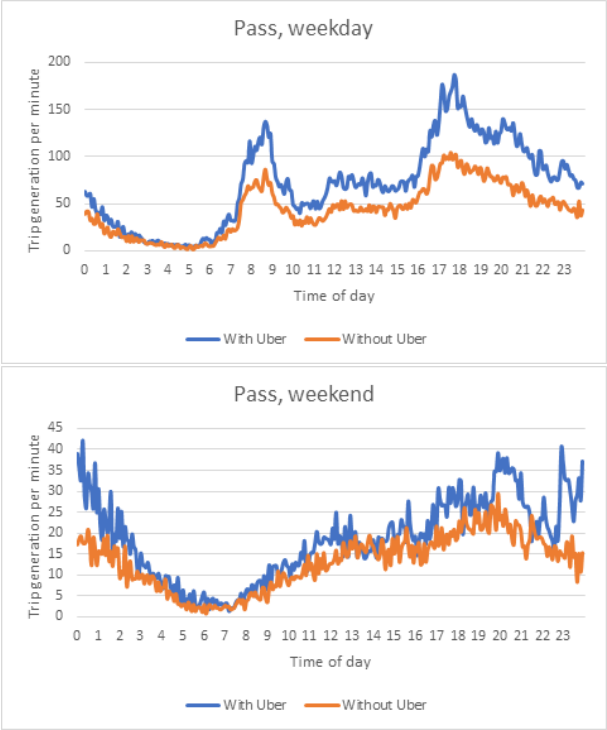


Figure 4.1: The distributions of temporal trip generation before and after Uber's exit (summer of 2016)

The figures reveal some interesting patterns. Pass-holders mainly use BSS on weekdays, especially during the morning and afternoon peak periods, which may be connected to commuting. Following the exit of Uber there is a significant decline in BSS usage, mainly during the commuting peaks. Changes in early morning and midday

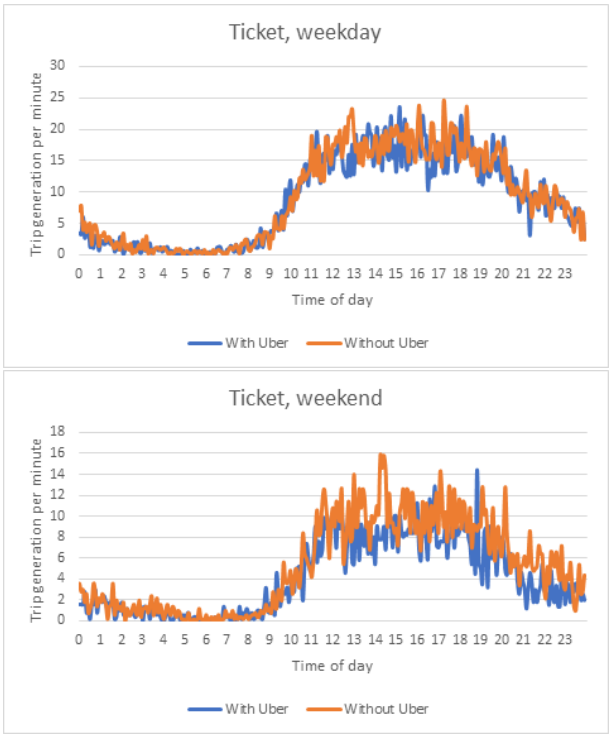


Figure 4.2: The distributions of temporal trip generation before and after Uber's exit (summer of 2016)

usage are less sizeable. Some usage reduction is also observable in evenings and late-nights. Furthermore, usage distributions indicate differences in usage during weekdays and weekends. While BSS is more frequently used during the commuting peak periods on weekdays, usage is the lowest during the morning period and it is mainly concentrated in the evening and night periods on weekends.

Ticket buyers' usage distribution is similar on weekdays and weekends, indicating that they might use BSS for purposes other than commuting. Their usage was altered significantly less than that of pass-holders' after the exit of Uber. There is barely any difference in the usage on weekdays and there is only a slight increase if any in weekend usage. However, as we have mentioned earlier only a minority of users use the service with tickets and the majority is using it with pass. Thus, the impact of the former subsample should be downweighted when considering the total impact of Uber's exit on BSS usage.

4.2 Regression results

The previous section revealed some interesting patterns regarding BSS usage. Yet, the changes in usage patterns might not solely be driven by the presence or absence of Uber, but be influenced by many other factors as well. As we have argued in the previous section weather conditions, network size and station-specific characteristics might impact the

usage of BSS, therefore a more thorough analysis in which we control for these variables is necessary to determine the impact of Uber. More specifically, a fixed effects panel model is estimated for the regression expressed in equation (3.1). Table 2.3 summarizes the estimation results.

Variable	**Pass**		**Ticket**	
	Weekday	**Weekend**	**Weekday**	**Weekend**
	(1)	(2)	(3)	(4)
Uber	1.742***	0.456	-0.404	-1.264***
	(0.569)	(0.655)	(0.251)	(0.375)
Network size	-0.035	-0.107	0.036	-0.054
	(0.056)	(0.075)	(0.025)	(0.043)
PET: Moderate Cold	-5.853***	-5.553***	-0.960***	-2.037***
	(0.562)	(0.788)	(0.250)	(0.451)
PET: Slight Cold	-0.971***	-1.665***	-0.340***	0.133
	(0.291)	(0.340)	(0.130)	(0.195)
PET: Moderate Heat	-5.050***	-0.950	-1.214***	0.181
	(0.425)	(0.638)	(0.189)	(0.365)
Precipitation: 0–5 mm	-3.356***	-1.926***	-0.346**	-1.431***
	(0.365)	(0.406)	(0.163)	(0.232)
Precipitation: >5 mm	-7.384***	-1.308**	-1.062***	-0.600
	(0.357)	(0.649)	(0.159)	(0.371)
Tuesday	4.056***		0.989***	
	(0.494)		(0.220)	
Wednesday	1.824***		0.901***	
	(0.593)		(0.263)	
Thursday	0.832		0.753***	
	(0.638)		(0.282)	
Friday	0.696		1.119***	
	(0.653)		(0.288)	
Sunday		0.657		-0.066
		(0.464)		(0.265)
N (sample size)	5,907	2,380	5,907	2,380
R^2	0.273	0.257	0.053	0.113

Notes: reference category for PET is No Stress, for precipitation is 0 mm and for the day of week dummies Monday and Saturday.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 2.3: Estimation results

As we have mentioned earlier pass-holders predominantly use the BSS on weekdays, while ticket-buyers use it more often on weekends (see Table 2.1). For this reason we concentrate our attention on the effects generated in these cases.

Estimation results for regular users (pass holders) are shown in the first two columns of Table 2.3. The first column of the table indicates that Uber had a positive effect on

BSS usage during weekdays. The results suggest that the market exit of Uber caused a decrease of around 1.74 trips on average per weekday per station. Considering that the average trip generation of a station on weekdays was 26.9 (see Table 2.2), this shows an approximate 6.5% decrease in trip generation. Given that there were 96 BSS stations in Budapest in the time frame considered, the exit of Uber *ceteris paribus* caused a decrease of around 167 rentings per weekday. These results suggest a complementary relationship between the two services.

The third and fourth columns of Table 2.3 show results for ad hoc users, who are using BSS with tickets. Results are exactly the opposite to the ones we observed for regular users. The presence of Uber had a significant negative effect on weekend usage. In numbers, the exit of Uber resulted in a 1.26 increase in average daily trip generation for a given station during weekends. This is rather substantial since it shows an approximate 23% increase is BSS usage. These results indicate that ad hoc users use the BSS as an alternative to Uber during weekends.

The negative effect of the network-size, even though is not significant, might be surprising. This might be the case because the popularity of BSS did not increase with the expansion of the network. Since the new stations were added on the outskirts of the city centre they were naturally less attractive, and less frequently used.

Not surprisingly, if thermal conditions deviate from the ideal, BSS usage generally decreases. Since no thermal stress is the reference category in the regression, the estimated parameters indicate that BSS usage is ceteris paribus lower for less favorable PET categories. For example, if there is a slight cold stress, BSS usage is lower by 0.97 trips per station per day on weekdays among pass holders. Similarly, BSS usage is lower by 5.05 trips per station per day on weekdays among pass holders if there is a moderate heat stress.

Precipitation is also negatively impacting BSS usage in general. Results indicate that a light rain during the day reduces BSS usage by 3.36 trips per station per day on weekdays among pass holders. However, there are some surprising results, namely that the effect of relatively light rain is stronger than that of a heavier rain during weekends. This can be caused by two effects. First, there were only 5 weekend days with precipitation above 5 mm. Additionally, we used daily averages, therefore, it is possible that it was raining at dawn or late night on these days, yet it did not affect BSS usage during the whole day that much.

The day of week dummies indicate that Tuesdays and Wednesdays became more important in BSS usage. Thursdays and Fridays do not differ from Mondays in the pass subsample, but there is a difference in the ticket subsample. The significant day of week variables in the ticket model indicate that ticket-based usage can change patterns across years. However, it is important to note that ticket buyers mainly use BSS on weekends,

therefore, the significant estimates rely on relatively small number of trips. Furthermore, there is no significant difference between Saturdays and Sundays based on the weekend model.

Table 2.4 shows summary statistics about station-specific fixed effects for the 96 stations in use. The station-specific fixed effects capture all the changes that occurred in the average usage of a station from 2015 to 2016 apart from the exit of Uber, thermal conditions and precipitation. The average values of the fixed-effects indicate that there is a slight increase in usage, mainly among pass holders. Usage with tickets did not change considerably.

Subsample	Average	Standard deviation	Min	Max
Pass – weekday	0.23	5.87	-15.38	21.69
Pass – weekend	0.16	5.75	-41.33	11.29
Ticket – weekday	-0.03	1.19	-3.88	5.76
Ticket – weekend	0.01	1.57	-3.17	8.02

Table 2.4: Summary statistics for station-specific fixed effects

The minimum and maximum values indicate that some pattern changes occurred in the analyzed period. Some stations gained popularity while others lost some. This might have been caused by the extension of the network in 2015 (15 new stations were opened in June 12, 2015 and 7 new stations in July 31, 2015). As the network size variable shows, the extension did not increase the overall usage of the system but affected the usage patterns of the existing users. However, the interquartile range of the fixed-effects is between -3 and $+3$ in all cases and the decrease was never above 50% of the average number of trips generated by a station during the summer of 2015 suggesting that the majority of the network was unaffected.

4.3 Regression results by time periods

Since the daily distribution of trips is uneven, we also investigated the effect of Uber in different time periods of the day. This method enabled us to capture the temporal differences in usage and shed light on how users combined Uber and BSS within a day. We identified five time periods: dawn (1:00-7:00), morning peak (7:00-10:00), midday (10:00-16:00), afternoon peak (16:00-20:00) and night (20:00-1:00) based on the usage distribution of BSS trips within a day (see Figure 4.1 and Figure 4.2) and on prior literature (see Faghih-Imani et al., 2017a and El-Assi et al., 2017).

Results are summarized in Table 2.5. Control variables were eliminated from the table to reduce its size. More detailed results are presented in the Appendix (see tables

Variable	Pass		Ticket	
	Weekday	**Weekend**	**Weekday**	**Weekend**
	(1)	(2)	(3)	(4)
Dawn	0.024	0.459*	0.018	-0.031
	(0.108)	(0.246)	(0.034)	(0.064)
Morning	0.132	-0.170*	-0.012	0.014
	(0.133)	(0.096)	(0.036)	(0.052)
Midday	-0.275	-0.358	-0.105	-0.547**
	(0.201)	(0.267)	(0.149)	(0.236)
Afternoon	1.298***	-0.112	-0.131	-0.395**
	(0.252)	(0.281)	(0.117)	(0.187)
Night	0.745**	0.686*	-0.120	-0.350***
	(0.290)	(0.364)	(0.091)	(0.135)

Notes: Fixed effect panel regression results (with an AR(1) error term in the weekday subsamples) using network size, PET scores and precipitation as control variables. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 2.5: Effect of Uber on BSS usage by time periods

?? to ??). The results shed light on the following patterns. For pass holders, Uber and BSS appear to be complements especially in the afternoon commuting periods on weekdays. The exit of Uber caused a significant reduction in BSS usage during the afternoon peak period and at night for these users. These findings support our conjecture that the presence of Uber might encourage commuters to leave their cars at home and use a combination of other transportation modes, including BSS, instead. For ticket buyers, Uber and BSS appear to be substitutes, and this relationship is statistically significant throughout the day except at dawn and morning. This appears to be convincing since a considerable share of the ticket users are tourists, who are likely to start their city tour later during the day and may use either BSS, Uber, taxi or public transport to travel within the city without having a plan to combine these transportation modes. If Uber is not available, BSS obviously will get a higher share. The more detailed results presented in the Appendix reveal somewhat counter-intuitive effects for the control variable in some cases. In particular, slight or moderate cold stress seem to have a positive effect on BSS usage especially in the afternoons and at nights. One can speculate that these thermal conditions might be even conducive to cycling on summer evenings.

4.4 Discussion

Policy discussions about Uber are quite widespread, with the sharing service generating a lot of controversy (for an overview, see e.g. Taylor, 2017). While other taxi service

providers and some critics charge Uber with "not playing by the rules" and "exploiting" their workers, many economists have emphasized the social welfare increasing effects of Uber: the ride-sharing service could increase competition in local taxi markets, providing cheaper and higher quality and quantity service. Uber also has a system of surge pricing that, according to studies conducted by Cramer and Krueger (2016), is effective in balancing supply and demand. There seem to be large direct benefits from allowing Uber to operate, particularly in terms of consumers' surplus (see Cohen et al., 2016), although some raise questions regarding the viability of Uber's business model and the long-term welfare effects (see e.g. Horan, 2017). All in all, open questions remain regarding the right regulatory framework with respect to Uber. However, less attention has been given to the policy implications of indirect effects caused by substitutory or complementary relationship between Uber and other local transportation services. Our findings suggest that many users of Uber have a preference for multimodal urban transport use. This means that if Uber is banned from a city, it might depress the use of other local transportation services, including BSS.[3]

Regarding bicycle-sharing, Fernandez-Heredia et al. (2016) found that convenience (flexibility and efficiency) and pro-bike attitudes have an impact on the demand for BSS. Thus faster and more flexible transportation (especially for commuting) seems to be a factor in BSS usage. Faghih-Imani et al. (2017a) showed that BSS is competitive in terms of travel duration with taxis in the inner city of New York. This is particularly true during peak times and for some specific routes where cars have to make a longer trip due to e.g. one-way roads or traffic jams. Not surprisingly, results by Campbell and Brakewood (2017) show that there is substitution between public bus usage and BSS usage in New York, however, the size of the substitution effect is rather small. These results indicate that BSS is an effective way of commuting in dense cities and could act as a substitute to cars. To adjudicate the policy implications from all this, we also need to take into account externalities from BSS use. *Ceteris paribus*, greater reliance on bicycles may contribute to decreased pollution levels (see Johansson et al., 2017) and furthermore, bicycle usage has numerous health benefits (although we do not count this among the externalities). In the light of our findings, therefore, we believe that expelling Uber might, apart from the first-order welfare effects, adversely affect some other policy goals. Naturally, the generalizability of our results obtained for Budapest is in question. However, we would again refer to the findings of Hall et al. (2017) who find a complementary relationship between public transport and Uber, suggesting that consumers prefer multimodal transportation in general. Thus, our results

[3]Hall et al. (2017) showed this to be true, under certain conditions, to traditional modes of public transport, while our study points to complementarities between bicycle-sharing and Uber.

might be relevant for several medium-large cities and these findings may provide a useful contribution to the debate on ride sharing services by empirically verifying and measuring the impact of Uber on BSS ridership on a medium sized city.

5 Conclusion

In the past few years, several innovations were introduced in local transportation. In this article, we analyzed the interaction between two new services, Uber and bicycle-sharing. BSS is getting more and more widespread and nowadays middle size and even small cities are setting up their own networks. Uber is currently present in more than 700 cities worldwide with an explicit aim to further expand its business and geographical footprint.

In this article we exploit the fact that Uber exited from the Budapest market after a regulatory change in the middle of 2016. This natural experiment makes it possible to estimate the impact of Uber on BSS ridership. Our results suggest that regular BSS users combine bicycle-sharing with Uber to commute, and, therefore, banning Uber caused an around 6.5% decrease in BSS usage on weekdays among regular users. On the other hand, *ad hoc* users mainly use BSS and Uber as substitute services, especially during weekends and the exit of Uber caused a 23% increase in BSS usage among these users on weekends.

The net effect of Uber on BSS thus depends on the usage frequencies of the two groups (regular and *ad hoc* users). Not surprisingly, the majority of the trips was generated by regular users, hence the exit of Uber caused an overall decrease in BSS usage. This finding draws attention to some unintended consequences that are worth taking into consideration in future policy decisions. Our results also provide valuable insights into commuters' preferences for combining different transportation modes, yet further research on this topic is needed.

Chapter 3

A Theory of Early and Late Specialization

1 Introduction [1]

This chapter examines incentives by parents to invest in their offspring's human capital in two possible directions: investments can be made in human capital complementary to specialized activities (such as musical training, sports or advanced technical knowledge), or human capital that can be complementary, perhaps to varying degrees, to many different activities (such as basic mathematical and reading skills as well as non-cognitive human capital). General human capital is mostly acquired in elementary and middle school, while college education is often the terrain of specialization. Liberal studies degrees, however, can be said to provide a fairly general stock of human capital. What determines if parents and their children will have a liberal arts education or choose a specialized field relatively early on? This area of study has generally been neglected but we believe it is becoming more important to study as it can shed light on the question of how technology-induced labor demand shocks influence investment in human capital. This essay primarily studies the effects of uncertainty and changes in uncertainty related to future job prospects. Early specialization can have advantages due to dynamic complementarities in the accumulation of human capital. Skills acquired later often build on (general and specific) skills acquired earlier.

This chapter attempts to explain the causes of some important developments in human capital investment patterns over the past few centuries. It also partially accounts

[1]The author benefited from conversations with Ed Lazear and Kevin Murphy on the subject. The author also thanks Barna Bakó, Tyler Cowen, Kevin Murphy and Glen Weyl for useful comments and suggestions, as well as several participants of the 2020 CUB Department of Economics workshop for their comments.

for heterogeneity in human capital investment patterns across households. An important development following the set of the industrial revolution was, and arguably, some of it has been caused by the greater availability of formal schooling, a *relative* decrease in early (childhood) accumulation of specialized human capital and a relative increase in investment in general human capital, such as basic numeracy, literacy, arts and sciences. At the same time, investments in specific human capital increases and typically occurs earlier in "superstar market". The shift has been usually explained with the wide availability of formal education. However, there is another angle which has not been explored in the literature but is nonetheless relevant: In modern economies the division of labor among individuals is largely organized on the principle of (formal) equal opportunity: each individual can freely decide which industry they would like to work in, and they will invest in human capital specific to the profession of their choice. For much of human history, however, this has not been the norm. Feudal and caste societies prescribed pre-determined "roles" for groups of people. When entry and exit across sectors is quite limited, workers may earn rents in their respective industry, while at the same time they might have forward knowledge on what sector they are going to have access to. This may be due to innate abilities, a family history of working in the given industry (that would increase parents' productivity in producing specific human capital, but also provide them with connections to industry "insiders") or restricted entry combined with privileged access to certain classes of people. All of this raises the returns to early specialization, leading parents to invest comparatively more in industry-specific, and comparatively less in industry-neutral human capital. Therefore, in such environments, young adults will often be proficient in one specialized task, while being fairly "illliterate" in "general knowledge". A specific example we will consider is the effect of the sexual division of labor on human capital acquisition and how this has changed over the centuries. When "gendered expectations" (either due to comparative advantage or simple "focality") coordinate a strict sexual division of labor, women invest little in market human capital, given that they are unlikely to work in the market sector. This may lead to an inefficient allocation of talent but at the same time makes human capital investments more productive. Most social scientists and commentators studying these phenomena emphasize that these pre-determines social roles were tools used by a "ruling class" or "the patriarchy" for its own benefit. This essay takes a more "functionalist" approach and ponders whether there have been some efficiency advantages to such arrangements and why they have declined over time. At the same time, we do not suggest that rent-seeking is not at least a partial explanation: we do suggest, however, that having beneficial efficiency properties certainly increase the likelihood of adopting such pre-determined social roles around specialization, while when they become inefficient, there arises powerful societal pressure against them.

Uncertainty in human capital investment decisions can be captured by assuming uncertainty over abilities (or endowments). This is not the route we take here, although arguably this is also an important channel through which specialization may be delayed. Instead of uncertainty over endowments we focus on uncertainty over *rents*. We do so because while learning one's abilities might also be costly, we believe that it is even more costly to know one's place in the ability *distribution* as this requires not only knowledge in one's abilities or other endowments but also the endowments of other individuals in the economy. This can be one source of uncertainty over rents.

Our results imply a potential tradeoff between talent allocation and investment in human capital: When individuals cannot earn large rents in an industry, apart from rents on natural talent, and/or which industry they are going to work in is not "arranged" ex ante talent is allocated more efficiently across the economy. However, consider a case when there are large entry barriers in an industry and certain individuals are pre-selected into working in the industry. At the most extreme, think of a prince or a princess pre-selected to be a king or a queen, or children expected to continue the trade of their parents. In such cases, there are large returns from human capital investment in the given industry. Pre-selected individuals thus might achieve greater productivity in the particular industry, and if general and specific human capital are complements, they will probably accumulate more human capital *in general*, while at the same time talent allocation is hurt as they could have chosen an industry where their natural talent is greater. When, however, there is no selection whatsoever, human capital investment is likewise efficient: the agent randomly chooses a sector to specialize in, and invests efficiently in specific and in general skills. The biggest problems arise in interim cases. There agents are induced by the possibility of rents in several sectors, which leads to either an excessive *relative* accumulation of general skills or "malinvestment", if certain groups of people can expect to be "privileged" with only some probability. Our discussion on this tradeoff has some similarities to the well-known tradeoff (in certain cases) between competition and investment incentives, only in our case the investment occurs in human capital, not innovation or physical capital. We will see, however, that "equal opportunity" is almost always first-best efficient, while a strict prescription of roles can be second-best efficient (serving as a local optimum) if there are already (loosely) established roles in the economy.

A paper that is close to (yet still different from) our approach is one by Lazear (2009), who studies investments in general and firm-specific human capital. The novelty of his approach is that he posits that all human capital is general and the different across skills specific to given industries is that different industries require different combinations of general skills. Although we take a different, and more traditional, approach, some of our results are closely related to Lazear's. Like Lazear, we find uncertainty a key factor

in deciding what kind of human capital (or human capital combination) to invest in.

Another strain of research close to our approach is the one dealing with the effects of (potential) technological shocks on human capital acquisition. A paper particularly close to our approach is a study by Deming and Noray (2020). Deming and Noray study the effects of the possibility of skill obsolescence on age-earnings profiles. Like us, they find a tradeoff between increased productivity from specific human capital investments and adaptability to changing circumstances that requires more general human capital. Unlike them, we do not focus on technological (or other) changes, although we can interpret our very general model as incorporating possible technological shocks. For instance, one might use our general framework to study the choices of adults with preexisting experience in the labor market who face uncertainty about future labor market prospects. However, an equally suitable interpretation of our model is that parents face uncertainty about which sector their child will earn the most in her adult life.

Gervais et al. (2008) also focus on firm-specific and general human capital and the role of uncertainty in human capital investment decisions. Their approach differs from ours in two ways: first, they do not consider industry or occupation specific human capital, and second, in their model agents make their human capital investment choices in one period: therefore the authors do not consider possible dynamic complementarities in human capital accumulation.

Finally, we are indebted to the literature on industry-specific human capital, starting mostly with Derek Neal (1995)'s contribution. Before Neal's paper most of the attention has been focused on firm-specificity and not so much on industry-specificity.

One possibly surprising result we arrive at is that when the agent cannot earn a rent in any sector, she may still specialize early: she will be indifferent of where to specialize so will choose randomly and that way will maximize her adult income. [2] What changes when there is a possibility of a rent? Nothing if there is only one sector where the adult may earn a (quasi-)rent (although, in general equilibrium this can look different, and we will return to this question later). However, when there are many potential rent-generating occupations, the picture is not so clear. For if she specializes in sector s_i, she may miss out on the rents that she could enjoy in some sector s_{-i}. The *possibility* of rents in another industry induces the parent or the individual to delay specialization somewhat and invest a higher share of her material and time budget on accumulating general knowledge and skills.

Our essay is also a contribution to the literature on the effects of barriers to entry. Entry barriers to different markets and other causes of economic rent have been studied in economics since at least Adam Smith. Little research have been done, however, on the

[2] For more details of this argument, see Becker and Murphy, 1992, as well as Jaffe et al., 2019.

effects of entry barriers and other sources of economic rents on human capital acquisition. Rents, as we conceptualize them, can be the result of differences in natural talent (someone being more talented in a task than others), "artificial" barriers to entry, or it might be the result of having "connections" to important individuals in the given industry. Arguably, in certain industries, such as entertainment, politics and even in some business sector, social networks are an important determinant of how likely one finds a job in a given industry and how much she will earn.

This chapter is a contribution to the literature on "superstar industries", first spelled out by Rosen (1981). As market size increases and new technologies emerge that make possible mass joint consumption, industries are increasingly dominated by fewer actors who are able to earn a rent. The possibility of such rents may induce a relatively early accumulation of specific human capital, but if one is not sure which industry she would likely earn a rent in, this induces delays in specialization. An implication of our model is that a large part of the observed differences between say sportspeople or actors come down not on innate abilities, but accumulated human capital, although such human capital accumulation is incentivized by real or perceived differences in ability.

Finally, we build on the literature on dynamic complementarities in human capital investments (as discussed e.g. by Cunha and Heckman, 2007). The accumulation of human capital depends on pre-existing skills ("skills beget skills"). In our model, the main insight we derive from this literature is that there can be (weaker or stronger) complementarity between earlier and later accumulated specific human capital. This is important because this creates incentives for investing in specific skills relatively early on. At the same time, early specialization increases the risk of "malinvestment", that is, human capital investments that will not be utilized by the adult child because she would be better off specializing in some other task.

The chapter is built as follows: in the first part we introduce a very general model and derive some simple results for specific and general human capital investments as well as total investment. In the next section we consider a limit case of our analysis: an economy with two sectors and a strict hereditary or "caste" system. Next, we consider a somewhat similar configuration: "traditional" gender roles under a sexual division of labor. Finally, we draw some policy-relevant conclusions and consider some extensions, including investment in industry "connections".

2 Specific and general human capital

As we have indicated, this essay is concerned with industry-specific and industry-neutral or general human capital. Two notes are worthy to stress here: the literature did not

focus on this issue as it has been assumed, rightly, for the most part, that everyone accumulates both general and industry-specific human capital, as the latter builds on the former. However, it should be stressed that basically all human capital investment builds on earlier accumulated human capital, whether general or specific. So investing in specific skills earlier can come with a high return, especially when the given sector does not require a lot of general skills. Sports, music or acting are good examples of this phenomenon.

3 The Model

Consider a model consisting of two periods. In period I. the parent of a child decides how much time to devote to investments of different types of human capital for her children. In the second period, the adult child decides which sector she enters as part of the labor force, and if she enters a profession, she has the opportunity to accumulate additional sector-specific skills. We assume that while the parent has imperfect knowledge over which sector the child would likely work in, the adult child has full information over the returns of human capital investments made in their adult period. In the economy, there is a set $s = (1, 2, ..., S)$ of activities a worker can specialize in which we treat as exogenous. Parents can invest in human capital that increase productivity only in a given specialized activity H_s or in human capital that increases productivity in all of the activities, $\underline{H}$. Let W_s denote the per unit "price" of human capital, that is equal to the marginal product of the worker, that can be earned in activity s. After investments by the parent, the adult child can invest further in her specific human capital. We notate the stock of adult "training" H_t. We make the following assumptions:$W_s'(H_s) > 0$, $W_s'(\underline{H}) > 0$, $W_s''(H_s) < 0$, $W_s''(\underline{H}) < 0$,$W_s'(H_t) > 0$ and $W_s''(H_t) < 0$. or the sake of simplicity, let $W_s = W_0 + \theta_s\mu_s$: W_0 is a "basic" rental price of human capital, while with some probability θ_s the adult child can earn a rent μ_s in industry s. That is, she can earn more in that industry than in other industries. Importantly, we assume that the adult child will earn a rent in only one industry: if this was not so, she would choose the industry with the highest rent, or, if rents are equal across industries, she would be indifferent among the industries. Uncertainty, and hence the problem we focus on, matters only to the extent rents in industry s and $-s$ are mutually exclusive. One can think of these rents as stemming from differences between individual talents. While parents are aware of the *absolute* level of their child's talent in a given area, they have imperfect knowledge about their offspring's position within the talent *distribution* in the relevant industry. It is thus uncertainty over "relative" abilities or other endowments, not uncertainty in absolute abilities (endowments) that we consider. A parent may be confident that her child would be a good doctor, yet she does not know how many equally good or better doctors there would be with whom her child would have to compete with. This assumption gives "teeth" to our theoretical framework. If uncertainty was over absolute abilities, we could not make any prediction on what families would specialize earlier or later. Of course, we could introduce uncertainty in both dimensions, however, it would complicate the model without any analytical benefit.

Note that as long as the probability distribution is smooth and the error terms of individual probabilities are zero, a decrease or increase in any θ_s changes the incentives only in choosing the level of specific human capital investment across activities, but

not across specific versus general human capital investments. This means that greater uncertainty does not effect incentives for specialization as long as the investment problem is convex and continuous.

These assumption, however, are quite unrealistic. In most real-life situations, when parents choose to invest ins specific or general skills for their children, they face a non-convex choice problem that is mainly due to the presence of fixed costs. Also, specialization involves increasing returns of scale (as we have mentioned above, and as shown by Becker and Murphy (1992)), but we choose to represent scale efficiencies simply by introducing the fixed cost of accumulating specific human capital. For that reason we include the following auxiliary assumptions:

$$\sum_{s\geq 2} W_s(H_s, \underline{H}) < \sum_{s\geq 2} f, \tag{3.1}$$

and

$$\max W_s(H_s, \underline{H}) > f, \tag{3.2}$$

where f is the fixed cost of investing in any kind of specific human capital, meaning that it is always worth investing in one, but at most one, activity. We assume that workers' utility in the second period is equal to their income. We do not explicitly consider non-monetary benefits and costs of working at a job, however our analysis would not change too much if we considered such non-monetary considerations. In fact, our formulas for wages could be interpreted as "full incomes", which could include the non-monetary benefits of a job.

The parent maximizes

$$\max_{\underline{H}, H_s} u(C_p) + a\delta \max\left\{(W_0 + \theta_1\mu_1)(H_1, H_t, \underline{H}); ...; (W_0 + \theta_s\mu_s)(H_S, \underline{H})\right\} \tag{3.3}$$

subject to the time constraint

$$T = C_p + \underline{h} + h_1 + ... + h_S + f \tag{3.4}$$

and the human capital production functions $\underline{H} = f(\underline{h})$ and $H_s = f(h_s)$ for every s, where $u(C_p)$ is the utility from the parent's own consumption, T is the "time budget" of the parent, W_0 is the "base" rental price of human capital, δ is the discount factor, μ_s is the rent earned in sector s, θ_s is the probability of the rent occurring in sector s, $\underline{h}$ and h_s are the time units spent on accumulating general and specific human capital, respectively, and a is the altruism parameter.

The first order conditions are

$$a\Bigg(\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s}+\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s}+ \\ \frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\Bigg)=\lambda \tag{3.5}$$

as well as

$$a\Bigg(\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial\underline{H}}\frac{d\underline{H}}{d\underline{h}}+\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial\underline{h}}\Bigg)=\lambda. \tag{3.6}$$

Both general and specific skills produce two effects: they raise productivity directly, as the first term in each FOC shows, but they also contribute to further skill development. Both general and specific skills contribute to later accumulated specific skills. From the first-order conditions, the following conditions are derived:

$$\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s}+\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s}+ \\ \frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}= \\ \frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial\underline{H}}\frac{d\underline{H}}{d\underline{h}}+\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial\underline{h}}, \tag{3.7}$$

$$\frac{\partial v(C_p)}{\partial C_p}=a\Bigg(\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s}+ \\ \frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s}+\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\Bigg) \tag{3.8}$$

and

$$\frac{\partial v(C_p)}{\partial C_p}=a\Bigg(\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial\underline{H}}\frac{d\underline{H}}{d\underline{h}}+\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial\underline{h}}\Bigg). \tag{3.9}$$

Equations (4.8) and (4.9) state that the marginal benefit of investing in the child's (general or specific) human capital must, in optimum, equal the marginal utility per dollar of spending on the parent's own consumption. As usual in human capital models, the more altruistic the parent is, the more she will invest in her child. Equation (4.7) is "new", relative to earlier literature on human capital. It expresses that the marginal return on investing in general and specific human capital must be equal in optimum.

Notice that later accumulated specific skills build on both specific and general skills.

If they are more complementary to earlier accumulated general skills, than an increase in uncertainty over future rents leads not only to more general skills accumulated, but also to more specific skills, as later skill development can substitute for earlier skill accumulation. More precisely, if early and later accumulated specific human capital are perfect substitutes, the parent in effect compares the marginal return to later and earlier accumulated skills as well as the marginal costs of acquiring it. The marginal rate of return to later accumulation can be expressed as $\max \sum_s \theta_s \frac{dH_s}{dh_s} \underline{H}$ as $\sum_s \theta_s$ is the probability of the adult child earning a rent in any industry. The marginal rate of return of earlier specific human capital accumulation is, on the other hand, $\max \theta_s \frac{dH_s}{dh_s} \underline{H}$. The marginal rate of substitution is thus $\frac{\sum_s \theta_s}{\max \theta_s}$. If the "price ratio" is not greater than 1, the parent will not invest early in specific skills, rather, she will leave this "task" to the adult child. It is, of course, not the parent who makes the decision to invest in specific training in the second period, but the child. If earlier and later accumulated specific human capital are on net substitutes, an increase in H_s will cause a decrease in the marginal return on H_t, while an increase in H_t decreases the marginal return on investing in H_s, given the usual concavity assumptions.

If later accumulated human capital builds more heavily on earlier accumulated specific skills, than such substitution is possible only to a very limited degree. In this case total investment in human capital can potentially be increased by decreasing uncertainty over rents: instead of the possibility of earning a moderate rent in multiple sectors, if the adult child can only earn a large rent in one sector, the parent will invest more in specific and general skills (due to complementarity between the two). However, if the child is expected to earn a high rent with some probability in different industries, the expected return on both general and specific human capital will be high, leading to a quite substantial level of total investment in human capital.

A question arises that apart from uncertainty over rents as well as the extent of rents influencing the relative amount of general and specific investment, how they influence total investment in human capital. Total investment increases when $\max \theta_s \mu_s$ when specific and general human capital are complements, while it does not change if they are on net substitutes. An increase in $\sum_s \theta_s \mu_s$ increases investment in general skills, which in turn increases total investment if general and specific skills are complementary. Otherwise, total investment is unchanged. Parents of children with many potential "relative" talents will, under complementarity between general and specific human capital, in general spend more on both general and specific education than parents of children with "low relative talents" (no possible rents in any industry), while parents of children with "one great talent" may also invest much in the case of complementarity. Notice also that just being a person with many (potential) talents does not guarantee rents in any industry. What

causes prospective rents to appear is the relative scarcity of their talent. We can can conjecture that at the high end of talent distribution talent is relatively scarce. On the other hand, pupils who are less talented in any areas do not have this patterns of scarce talents, therefore, they cannot expect to earn rents or quasi-rents in any industry.

The incentive for parents of children of many talents to invest relatively more in general education stems from two similar but still distinct consideration: first, their adult child can directly use the accumulated general knowledge and skills in many industries. Second, they can build on general skills later in the accumulation of specific skills in many industries. In contrast, early specialization results in the preponderance of skills that can be built on in only one industry.

What happens, when the adult child cannot expect a rent in any industry? Then, she earns the same wage everywhere ($\mu_s = 0$ for any s). How does the parent decide which sector to invest in for her child? As she is indifferent between the sectors, she will pick a sector randomly and invest in specific human capital there. She will not invest that much in general human capital *in relative terms* as she does not need to "insure" the child against the "risk" of being better at a task different from what the parents invested in. In a sense, and perhaps surprisingly, the lack of any potential rent is the case with the most complete early specialization (even though, in *absolute* terms, parents of would-be "superstars" invest more in specific human capital). To see this more clearly, consider that the optimality condition for such a parent is

$$\frac{\partial W_0(H_s, H_t, \underline{H})}{\partial H_s}\frac{dH_s}{dh_s} + \frac{\partial W_0(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s} + \frac{\partial W_0(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s} = \frac{\partial W_0(H_s, H_t, \underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}} + \frac{\partial W_0(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}. \tag{3.10}$$

If the marginal products of different human capital investments are the same, then essentially the lhs and the rhs is equal. In the same case, the rhs in the case of positive potential rents would be greater than the rhs, and equality would be satisfied only if the marginal product of investing in general human capital is smaller: this comes about if relatively more is invested in general skills. But in the no-rents-case, the two sides are equal when marginal products are equal. That is, ceteris paribus, a parent who expects no rent for her child in the labor market will invest *relatively* more in specific human capital early on.

Notice that while low-ability individuals will specialize early, they (although this is not captured explicitly in the model) will not specialize very intensively as often they have a lower learning ability. They will invest little in both general and specific human capital, but the share of specific capital in their portfolio is typically large.

A further comparative statics result we can derive is the effect of μ_s on general and specific investment. An increase in the rent in a given sector increases specific human capital investment if and only if , it occurs in the industry that already promises the highest rent, or if the change results in an industry promising the highest rent. . On the other hand, it always leads to an absolute increase in general human capital investment. However, if the rent increase occurs in the industry for which the parent prepares the child with specific investments, the increase in specific investment is *relatively* greater than the increase in general investment, even if it is the same in absolute value. The ratio of specific to general investment is thus increases when the adult child can earn a bigger rent in the industry her parent invests in.

In certain occupations, such as acting, modeling and sports, individuals not only accumulate specific skills early on, but also start working in the industry relatively early on, often during childhood. The demand for child actors etc. might provide an alternative explanation for early specialization to the explanation we provide, however, it is just as likely that workers start working in these industry at an early age *because* they already accumulated substantial specific human capital. Another example we can bring up is chess: talent that is complementary to playing chess is often discovered early on, and later chess skills heavily build on earlier accumulated chess skills.

Whether parents of would-be "superstars" also increase their absolute level of investment in general human capital depends on the extent of complementarity between it a specific human capital. Acting or sports may require relatively little general (such as mathematical) skills, so families with children with such prospective talents will often make their children "skip" standard schooling. Likewise, certain "low-level" jobs also require relatively little general skills but a large amount of specific skills. These children will tend to "skip" grammar school or college and go to vocational schools. Which "class" of people will spend more time on investing in human capital *in general* is not so clear. Parents of would-be superstars spend a lot on specific training while perhaps invest little in general skills. Likewise, parents of children "at the lower end" spend a great deal on specific skills, while to some extent "skip" the accumulation of more general skills. It is likely parents of "children of many talents" who end up spending the most time with human capital accumulation. They invest a lot in general human capital and if that is complementary to specific skills, they will also spend a lot on those skills.

3.1 The effect of market size and technology

The industries which are superstar markets today have not always been that way. Professional football or basketball, for instance, are relatively recent phenomena. "Pop stars" and other

superstars were also a much rarer phenomena centuries ago. Id did not pay to prepare a child early for, say, a career in sports or acting first, as the markets fro these services were a lot smaller than the global markets they command today, and second, because production technology was vastly different. New technologies made it possible to serve larger markets, contributing to the emergence of superstar industries. This corresponds to a higher μ (as shown by Rosen, 1981), while at the same time it does not imply a lower θ_s, as it does not influence the talent distribution. Moreover, parents of especially talented children were, before the advent of mass media consumption, less able to compare the (raw) talent of their child against the demands of an industry and the talents displayed by its producers. This could either have corresponded to a greater or to a lower θ_s for any industry, however, Lazear (2016)'s statistical theory of overconfidence in occupational choice suggests that receiving more signals (through e.g. mass media) might increase overconfidence in one's (or one's child's) abilities relative to others as these signals may be noisy. Ex ate unbiased signals often lead to ex post overestimation of talent. The general increase in market size and the emergence of new technologies could have affected many μ_s values. This, ceteris paribus, means that when markets are bigger and due to the emergence of new technologies it becomes easier to serve large markets, we see both more early specialization of the "superstar" kind, but, as for many families uncertainty over rents is bigger, also more families sending their children to general middle schools as well as liberal arts colleges. Meanwhile, we do not see a corresponding increase in training for "craftsmanship", as families with relatively low expected future rents in any sector do not invest as much more in their children's human capital than if the child would have "brighter" prospects in a large market.

As we have suggested earlier, more and better information due to technological progress in general makes human capital investment more "efficient": parents of the most talented children can spot the relative superiority of their children's abilities and hence they invest earlier and more intensively in specific human capital, and importantly, they do so in the sector where their offspring actually has a comparative advantage.

Another thing to consider is the effect of an increase in the general income level. As in Rosen (1981), an increase in consumers' income raises the shadow price of consuming a large quantity of time-intensive household commodities, making consumers more willing to substitute quality for quantity. This further increases potential rents and amplifies the effects we have described above.

A further implication of our model is that superstars, to a large extent, are not born, but made. In particular, the larger is the potential market for a superstar's services, the "better" the superstars will be at what they are doing as a larger market size induces earlier and more intensive specific skill acquisition. This also implies that a large market

size leads to the very top talents becoming superstars at earlier ages (due to their already substantial specific human capital), while many in the same profession "blooming" later. Some causal observation seems to be consistent with this prediction of the model. For instance, Bobby Fischer became the youngest ever chess grandmaster at 18. Judit Polgar bettered this record with becoming a grandmaster at age 15. Current record holder Magnus Carlsen became a grandmaster at 13.

3.2 The role of learning-by-doing

If specific human capital is built primarily through education or training then early specialization at least partially crowds out later investments: earlier investment increases wages in the sector in which the skills are accumulated, which raises the opportunity cost of further learning. Returns to scale/dynamic complementarity may well offset this effect, but the main point remains: there is less of a positive relationship between early and late specialization if the main channel of specialization is schooling or on-the-job training. However, when the human capital accumulation happens mainly through learning-by-doing, then specializing early leads to both more labor supplied in the specific sector *and* further accumulation of specific human capital. The predictions of our model seem to be borne out by common observation. Actors, sports players or chess players often start working in the industry quite early, presumably not only because this type of work can be done from a relatively early age, but also because much of the learning is learning-by-doing.

3.3 An increase in general knowledge

Specialization builds on the accumulated general knowledge of humankind (Becker and Murphy, 1992). Scientific developments increase the return on both general and specific investments. The latter is true because specific knowledge builds on general knowledge, and the more advanced general knowledge is, the more individuals can build on in their specialized investments. For instance, an increase in general medical knowledge makes surgeons more productive through better diagnoses and better technology. Improvements in general knowledge can exacerbate the inequality discussed in the previous section, as students who are less able to absorb general knowledge will, due to the complementarity between general and specific knowledge, also invest less (although relatively more) in specific human capital.

3.4 Technology shocks

As we aim to provide a fairly general model of early and late specialization, a natural question that arises is how well our framework can accommodate the effects of technological change that makes some jobs obsolete and increases demand for other specific skills. We cannot do so in a straightforward manner as we only have two periods in the model. However, θ_s may be interpreted as the probability that a technology shock causes rents to emerge in the given industry, while μ_s can encapsulate, apart from the magnitude of the rent, also the duration of the rent. If rents are short-lived, this corresponds to a low μ. The predictions of the model then extend to the case of uncertainty due to technological change. For instance, an increase is $\frac{\max \theta_s \mu_s}{\sum_S \theta_s \mu_s}$ corresponds with relatively more and more intensive early specialization while a decrease in it corresponds to more delayed specialization. Our framework, however, is, in its current form, more difficult to apply to cases where technological shocks cause not only moderate shifts in demand but complete creation and eradication of industries. Then, with some probability, the adult child would end up with completely obsolete specific skills. In such a model it would still be true that more uncertainty induces delayed specialization. However, obsolescence of skills is arguably a bigger threat to lower skilled workers: so unlike in our general model, uncertainty due to potential obsolescence would be greater among the lowest ability students. An implication of this could be that under large degrees of technological uncertainty, even low ability students can gain from delaying specialization. More formally, consider the following modification of our model: let the expected wage in sector s be

$$W_s = (W_0 + \theta_s \mu_s)(\underline{H}, H_t, H_s)(1 - \Delta_s), \tag{3.11}$$

where Δ_s can be interpreted either as a skill obsolescence rate or the probability that the specific skill becomes completely obsolete (the industry disappears). Now instead of S existing industries, assume there are S potential industries. Assume further that as one industry contracts or disappears, another industry expands or emerges, hence $\sum_S \Delta_s = 0$. The first order conditions can be written as

$$a(1-\Delta_s)\left(\frac{\partial(W_0 + \max \theta_s \mu_s)(H_s, H_t, \underline{H})}{\partial H_s}\frac{dH_s}{dh_s} + \frac{\partial(W_0 + \max \theta_s \mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s} + \frac{\partial(W_0 + \max \theta_s \mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\right) = \lambda \tag{3.12}$$

as well as

$$a\left(\frac{\partial(W_0 + \sum_S \theta_s \mu_s)(H_s, H_t, \underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}} + \frac{\partial(W_0 + \sum_S \theta_s \mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}\right) = \lambda. \tag{3.13}$$

The threat of obsolescence does not influence directly the return of investing in general skills, but lowers the return on any specific skill. For that reason the prospect of technological obsolescence causes a change in the optimal human capital portfolio: it will consist relatively more general and relatively less specific skills. We did not assume anything about the determinant of Δ_s, but it is possible that it is larger in industries which require a lower level of general knowledge and where manpower can be more easily substituted for machinery. If this is so, then "weaker students" on the one hand can gain relatively more by investing more in general skills, yet often they have worse learning technology and they are also indifferent among sectors, with the latter two pushing them toward early specialization. That being said, obsolescence does cause every student to invest more in general human capital.

In building our model, we postulated that individual children may have different "abilities" in learning: some children learn more efficiently than others. Furthermore, learning general skills might command different abilities than learning general skills and especially general knowledge. "Good students", in particular, are better at accumulating general knowledge, while "poor" and "average" students have less of such abilities. Hence, they will spend less time accumulating general human capital and specialize earlier.

All of this has important implications for responses to technological disruptions. When a disruption occurs and some skills become (largely) obsolete, individuals who completed grammar schools and/or liberal arts degrees will have less difficulty in adjusting to the disruption, first because some of their skills can be put to direct use in any industry, and second, because of the "learning to learn" effect.. Those, on the other hand, who specialized early will face great difficulties. This implies that students who would have otherwise specialized early would greatly benefit from a delay in specialization. However, many of them specialized early in the first place because they are not "good learners". Hence, some of them will fall behind in their education.

Note that instead of "ability", we could also speak about "endowments", and it contains not only innate abilities but also capabilities acquired in early childhood (see e.g. Cunha and Heckman, 2007). Children from disadvantaged backgrounds (such as those born to single mothers) can greatly benefit from early childhood education programs all the more when uncertainty is greater and thus general education becomes more important.

3.5 Other demand shocks

Rents or quasi-rents may appear in an industry due to temporary or permanent demand shocks not necessarily connected to large technological changes. For instance, as incomes rise, consumers may shift their consumption toward more expensive goods and services.

Our framework can also make sense of the effects of such demand shocks. Rents, in particular rents that are not overly short-lived, can be created as a result of demand shocks if there is a relative scarcity of talents in the industry experiencing the positive demand shock. In our framework, θ_s may be the probability of a positive demand shock. As rents will still be the result of scarcity of talent or opportunities, it would still be the case that general good students will specialize later and would-be superstars and "bad students" would specialize earlier.

4 Discussion

Our model sheds light on an interesting tradeoff: larger and less uncertain rents create more incentive for human capital investment, especially investment in specific, but also in general human capital. However, with early specialization there is a greater risk of talent misallocation: later on it might turn out that the worker would have been more productive in another task. A good example of this tradeoff might be hereditary positions such as the position of a monarch. Monarchs invest heavily in a highly specific form of human capital for their children so that they can be good future monarchs (in the hereditary monarchs' case, $\theta_s = 1$). At the same time, hereditary positions usually lead to talent misallocation: the ruler who establishes a dynasty might be very able, but his or her successors less so, the simple reason being reversion to the mean. Such reversion to the mean occurs in dynasties in general. A royal dynasty, although not a good practical motivating example, can nevertheless be regarded as the pure form of such occupational dynasties. We arrive at a quite different conclusion when we consider a decrease in uncertainty due to better information on future potential "natural" rents on scarce talent. Better information about the scarcity of one's talents, perhaps due to technological changes that lower information costs, *increases* the efficiency of both talent allocation *and* the efficiency of human capital investment. If a parent knows with a high probability that her child will be good compared to others in say singing or chess, she will choose to invest early in specific skills, and this leads to better talent allocation given that the child is more likely to accumulate human capital in the sector when she has the highest comparative advantage.

To see this tradeoff more clearly, consider an economy with two persons and two industries, with both persons enjoying a "privilege" serving one industry. Both parties will invest heavily in specific human capital as they can be sure of where they are going to work. This would mean ceteris paribus higher wages in both sectors. However, as talent is drawn randomly, they will not necessarily specialize according to their comparative advantage. If ability is not uncertain, limiting entry into a profession is never optimal, as individuals simply, and optimally, specialize according to their talents. But both uncertainty about ability (not explicitly modeled here), and uncertainty about in which industries the individual would earn a rent decreases specific and increases general investments.

5 "Castes" and hereditary positions

For a large part of human history, occupations were "hereditary" in nature. Sons typically followed fathers in their occupations. Countries were ruled based on the hereditary principle. And certain occupations were available only for certain groups or "castes", access

typically determined based on whether someone had been born to the right class or caste. Becker (1981/1991) has a short discussion of such phenomena. He emphasizes that for many centuries families were the main institutions for the education of youth. That was an era where the economy was relatively static, so the young could expect to find work in the same sector as the family's elders. They accumulated specific human capital early and that contributed to the hereditary nature of occupational choice. One puzzle we believe is left to be answered is why occupational choice was often restricted to the hereditary principle by law or custom. Here we posit a hypothesis to answer this question: we suggest that occupational caste systems have developed partly to increase early investments in specialized human capital, by decreasing uncertainty about which industry any particular individual will end up working in. Our hypothesis may not be complete in that other forces presumably influenced these laws and customs. We believe, however, that our explanation is at least partially accounts for these phenomena. We consider a special case of our model in which economic mobility is severely restricted. In essence, certain occupations might be available only to certain segments of society. Alternatively, it could be easier for certain members of society to enter into certain professions due to connections to other workers in that industry. In our discussion we first analyze the case when one group enjoys exclusive access to a labor market (and another group may enjoy exclusive access to an other sector) and then extend our discussion to cases where individuals have some probability of being in a "privileged" position in an already protected market. The intuition we build is as follows: those who are in a "privileged position", that is, they are able to enter a protected profession, may enjoy rents in that profession, or at the very least, they can enter said profession with probability 1. For that reason they will specialize fairly early on and accumulate more sector-specific human capital, while perhaps also accumulating more general human capital, if the two are complements.

On the other side, those who are excluded from the sector will navigate into the remaining sectors, increasing labor supply there. This in turn may suppress wages in most sectors. A lower wage will induce lower investment in human capital (both specific and general!) among the group that is discriminated against.

5.1 "Separate but equal"

This subsection considers a caste system where there is no obvious hierarchical relation between the castes. We do so for the purpose of shedding light on a once-useful function of the hereditary principle: by decreasing uncertainty about who will end up in which profession, a caste system can coordinate specific human capital investments. Consider first a very small-scale society, one consisting of a village. In the village everyone knows

everyone else. Although individuals and households cannot observe each other's abilities and other endowments, they can observe human capital investments and they know everyone's occupation. In such a society, an equilibrium in which everyone follows his parent in the parent's profession will easily arise due to the "focality" of such an equilibrium and also due to parents or extended families being responsible for the education of children. Now consider an extended economy, with N individuals (or households). Now individuals cannot observe the occupation of each other, hence parents' occupation loses its focality. Suppose that $\frac{N}{2}$ individuals have a comparative advantage in task A, while the other half of the population have a comparative advantage in B. This means that one half of the population can earn a rent in A, while the other half in B sector, that is, they can earn a higher wage in one sector than in the other one. Assume further that market demand is distributed across the two sectors so that the market is in equilibrium when half of the population works in one, and the other half in the other sector. Finally, suppose that the population consists of two equally large demographic groups, which have the same talent distributions. This implies that the probability that any individual will end up working in the given sector is $\frac{1}{2}$. Now consider incentives to invest in general and specific human capital under a completely free market and under a "perfect" caste system. In a free economy, members of both demographic groups can work in both sectors. Because half of the population have a comparative advantage in task A, and the other half in task B, the equilibrium is straightforward: the allocation of labor occurs according to comparative advantage. Talent allocation is optimal but human capital investments may not be efficient. Consider the typical worker's parent's investment problem. Under a free market she maximizes

$$u(C_p) + a\Big(\frac{1}{2}(W_0 + \mu_A)(H_A, H_t, \underline{H}) + \frac{1}{2}(W_0 + \mu_B)(H_B, H_t, \underline{H}, H_t)\Big). \tag{5.1}$$

The first order conditions are

$$a\frac{1}{2}\Bigg(\frac{\partial(W_0+\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s} + \frac{\partial(W_0+\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s} + \frac{\partial(W_0+\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\Bigg) = \lambda \tag{5.2}$$

for specific skills, as well as

$$a\left(\frac{\partial(W_0+\mu_s)(H_s,H_t,\underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}} + \frac{\partial(W_0+\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}\right) = \lambda \tag{5.3}$$

for general skills. Under a caste system, when one demographic group is "assigned" to sector A, while the other is assigned to sector B, the first order conditions are

$$a\Bigg(\frac{\partial(W_0+\frac{1}{2}\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s}+\frac{\partial(W_0+\frac{1}{2}\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s}+ \\ \frac{\partial(W_0+\frac{1}{2}\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\Bigg)=\lambda \tag{5.4}$$

for specific human capital, as well as

$$a\left(\frac{\partial(W_0+\frac{1}{2}\mu_s)(H_s,H_t,\underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}}+\frac{\partial(W_0+\frac{1}{2}\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}\right)=\lambda \tag{5.5}$$

for general human capital. The intuition for these expressions is the following: under a free market system, both markets are served by those who are "the best". Everyone earns a "rent", but faces an uncertainty on which sector the rent will be earned in. This leads to a lower relative investment in specific human capital (and may also lead to "malinvestment" - investments made in one sector which then would not be used because the worker ends up working in an other sector: it would be more difficult to model this in our simple model, hence we omit discussing this issue). On the other hand, the payoff to invest in general human capital is greater than in the case of a caste system. This is because under a caste system, those with a comparative advantage in the assigned sector are grouped together with those who have a comparative disadvantage in it, hence, any worker will earn a rent only with probability $\frac{1}{2}$. In a free economy, however, as we have seen above, *everyone* earns a rent with probability 1. Another interesting property of our expressions is the relative size of the "base wage" and the rent. A base wage that is high compared to the rent means an individual can lose a lot by investing in specific skills that she ends up not using. A relatively high μ_s on the other hand means that the typical individual can lose more from a strict caste system, given that under a free market regime she would earn rents with certainty, while under a caste system she earn a rent with only a probability of $\frac{1}{2}$. Based on the above analysis we can derive a number of predictions as to when and why caste systems come to an end. First, a greater importance of general knowledge and hence general human capital makes the loss from a lower investment in general skills under a caste system more apparent and more pronounced. Second, the gains from a caste system are lower as learning-by-doing becomes less important *relative* to formal education, as then the self-productivity of specific human capital becomes weaker. Next, technological progress also should undermine caste systems. Unlike when uncertainty over labor market rents arises due to uncertainty over talent distribution and the difficulty of coordinating specific investments, when the uncertainty is due to technological change, everyone has

an incentive to make herself more "flexible" and less tied to a particular occupation in order to better adapt to changing market demand. Furthermore, an increase in mmarket demand in at least one sector can increase the rents of those who have a "true" comparative advantage in that sector and hence their incentive to demand an end to a caste system that restricts the sector's size. Finally, better information about one's place in the talent distribution oviously lowers the benefits, and increases the costs of occupational castes: parents realizing their child ihas a comparative advantage in sector A but she is assigned to sector B form pressure groups to change the system.

5.2 Hierarchical castes

In the previous subsection we considered a case when the neither caste is "privileged" over the other, but all is assigned to a specific sector. In reality, however, caste systems often reflected a social hierarchy, which was often thought to be "natural". Here we discuss the effects of a caste system where one caste is more privileged and the other and entry is limited to one of the two sectors, in one where only members of the higher caste can enter into. More formally, consider the case of members of two "castes", with members of a higher caste having exclusive rights to engage in sector I. This leads the rest of the population to shift their labor supply to sector II. The whole economy consists of these two sectors. Production in the two sectors depends entirely on labor input augmented by human capital. We assume that demand and supply conditions are such that originally wages are equal in the two sector, but due to perhaps some legislative change the supply curve in industry I becomes steeper above a certain quantity level, due to increased barriers to entry. Importantly, unlike in the previous subsection, we abstract away from "natural" rents: we assume rents are created solely by the caste system. We write down a fairly general model. We assume human capital production functions $H_s(, a_s, h_s)$ where a_s is ability related to accumulating human capital specific to sector s as well as $\underline{H}(\underline{h})$ for general human capital. The adult child can supplement her earlier education with "training", which can be characterized by the production function $H_t(a_s, h_t, h_s, \underline{h})$. The wage equations in sector I is given as

$$W_I = (W_0 + \mu_I)(H_t, H_I, \underline{H}), \tag{5.6}$$

while in sector II it can be written as

$$W_{II} = (W_0 - \mu_{II})(H_t, H_{II}, \underline{H}). \tag{5.7}$$

The first order conditions are

$$a(W_0+\mu_I)\frac{\partial H_I}{\partial h_I}\frac{\partial H_I}{\partial a_I}\left(\frac{\partial W_I}{\partial H_I}+\frac{\partial W_I}{\partial H_t}+\frac{\partial W_I}{\partial H_t}\frac{\partial H_t}{\partial H_I}\right)=\lambda, \tag{5.8}$$

$$a(W_0+\mu_I)\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_I}{\partial \underline{H}}+\frac{\partial W_I}{\partial H_t}\right)=\lambda, \tag{5.9}$$

$$a(W_0-\mu_{II})\frac{\partial H_{II}}{\partial h_{II}}\frac{\partial H_{II}}{\partial a_{II}}\left(\frac{\partial W_{II}}{\partial H_{II}}+\frac{\partial W_{II}}{\partial H_t}\frac{\partial W_{II}}{\partial H_t}\frac{\partial H_t}{\partial H_{II}}\right)=\lambda, \tag{5.10}$$

and

$$a(W_0-\mu_{II})\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_{II}}{\partial \underline{H}}+\frac{\partial W_{II}}{\partial H_t}\right)=\lambda. \tag{5.11}$$

Combining equations (16) and (17) we obtain

$$\frac{\partial H_I}{\partial h_I}\frac{\partial H_I}{\partial a_I}\left(\frac{\partial W_I}{\partial H_I}+\frac{\partial W_I}{\partial H_t}+\frac{\partial W_I}{\partial H_t}\frac{\partial H_t}{\partial H_I}\right)=\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_I}{\partial \underline{H}}+\frac{\partial W_I}{\partial H_t}\right) \tag{5.12}$$

as well as

$$\frac{\partial H_{II}}{\partial h_{II}}\frac{\partial H_{II}}{\partial a_{II}}\left(\frac{\partial W_{II}}{\partial H_{II}}+\frac{\partial W_{II}}{\partial H_t}\frac{\partial W_{II}}{\partial H_t}\frac{\partial H_t}{\partial H_{II}}\right)=\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_{II}}{\partial \underline{H}}+\frac{\partial W_{II}}{\partial H_t}\right). \tag{5.13}$$

Guaranteed access to the protected sector induces more early specialization if earlier and later specific skills are complements, and increases total investment in human capital if general and specific human capital are complements. The allocation of investment across general and specific human capital is "efficient" given perfect information, however, under an "equal opportunity" regime, with no barriers to entry, it would also be efficient as parents would simply invest according to their children's natural talent (we assume here, unlike in section I, that talent is not particularly scarce, so there are no natural rents). Talent allocation is efficient in this latter case and is inefficient under any caste system.

Who belongs to a privileged "class" is not always clear. Consider a less extreme version of our model: as above, there are two sectors (I and II), with barrier to entry in sector I. However, now there is no clear "caste system", rather, everyone can belong to the "favored" group with some probability. This will be relevant for the choice of investing in specific human capital. As before, individuals only invest in human capital specific to one sector. Under a full hereditary system, members of the favored group invest in human capital specific to sector I, while members of the disfavored group invests in sector II specific human capital. The situation is different in an imperfect system of favoritism. When an individual does not know if she will be favored or disfavored, she might invest in human capital specific to any of the two sectors, depending on the *expected* wages earned

in sector I and sector II. This opens the door to "malinvestment". The individual might believe, with some confidence, that she is favored and invest in sector I specific human capital, however, end up being disfavored an working in sector II. Thus, she accumulates human capital that she does not use afterwards. More formally, we can reformulate the above maximization problem by introducing a probability of being in a favored position θ, so that the expected wage in sector I becomes $\theta(W_0 + \mu_I)(H_t, H_I, \underline{H})$ and the expected wage in sector II becomes $(1-\theta)(W_0 - \mu_{II})(H_t, H_{II}, \underline{H})$. If the former is greater than the latter, even if the probability of being privileged is not high, the agent will invest in human capital specific to sector I. This is a further source of misallocation of resources, as if θ is relatively low while it is still worthy for the agent to invest in human capital specific for sector I, the specific human capital accumulated would likely not be used.

There is a potential tradeoff between investment efficiency and talent allocation. Under "equal opportunity", parents invest in their children according to their natural talents (for the sake of simplicity we assumed that there is no uncertainty over ability, or, if they have equal talent in each sector, the parents choose a sector randomly. Both talent allocation and investment are efficient. However, consider any situation where there is some non-meritocratic "assignment" of individuals to different sectors. Talent allocation is, to some extent, already hurt. If the assignment becomes stricter, it further aggravate the talent allocation inefficiency, however, it also leads those whose (children's) talents have already been "misallocated" to invest more in their human capital as they know more certainly which sector their child will work in. Note that the marginal cost of misallocation decreases in the "strictness" of assignment: if there is already a strong enough level of ex-ante assignment into professions, fewer potential labor will be newly misallocated. At the same time, the marginal benefit of further tightening of roles is also diminishing. If, however, the former decreases faster than the latter, the marginal benefit curve will intersect the marginal cost curve from below.[3] If θ goes above this point, then the marginal benefit will exceed marginal cost, hence, society will converge toward a strict hereditary system as long as there are societal pressures toward efficiency. If, however, θ goes below the intersection, societal pressures will likely be created to move toward equal opportunity, that is, the first-best efficient outcome. There is one more reason why a strict system of social roles might be second-best: in an intermediate case, there is substantial "malinvestment" in human capital, in that families are "lured" into making investments specific to the more lucrative sector, yet many of these families will end up being shut out of that sector, making their investment a waste. Our reasoning is illustrated in Figure

[3]More formally, the marginal benefit of an increase in θ can be expressed as $(W_0 + \mu_s)(H_t, H_s, \underline{H}) + \frac{dH_s}{d\theta} + \frac{dH}{d\theta}$ which is decreasing in θ given concavity assumptions. The marginal cost on the other hand is $\gamma(\theta)(H_s(a_s{}')\underline{H} - H_s(a_s))\underline{H}$, where γ is the fraction of individuals newly "assigned" to one of the sectors, with $\gamma'(\theta) < 0$.

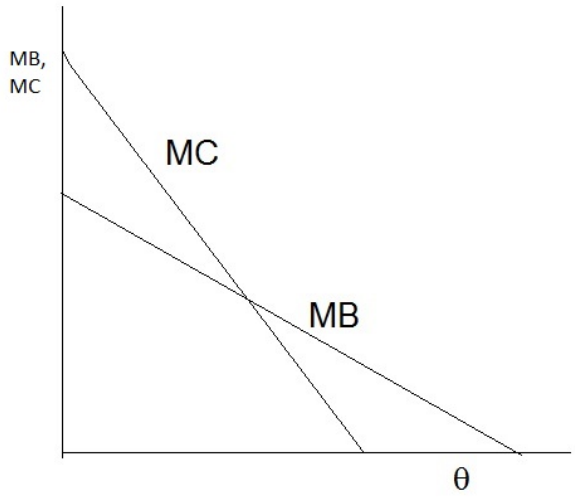


Figure 5.1: Marginal benefit and cost of "tightening the caste system"

3.1.

To put our main points in yet another simple way: barriers to entry obviously reduce efficiency in various ways, however, when roles are pre-determined, that is, individuals know which sector they will likely enter into, human capital investment can at least be second-best efficient. Talent-allocation is, however, greater with pre-determined roles. Hence, we may have an investment efficiency-allocative efficiency tradeoff. Importantly though, full "equal opportunity" is first-best: parents invest in their children's human capital according to their innate abilities or at random, hence there is no misallocation of talent and at the same time investment decisions are efficient.

The mechanism we study shares some similarities with patent races (see e.g. the seminal work by Loury (1979). In a patent race, would-be investors compete for being the first to introduce an innovation that then can be patented after which they can enjoy a monopoly rent. The patent race leads to what is called investment duplication: all firms invest in innovation, but consumers are indifferent over which firm should "win" the race, therefore, some of the investment by firms leads to social waste. A key difference between patent races and our case is that we do not model a race, however, just like in a patent race model, investments (in our case: investments in human capital) can represent a waste.

To sum up, a strict hereditary or caste system that creates rents lowers uncertainty over rents, increases total human capital investment in the restricted sector, but causes both underinvestment in human capital in the unrestricted sector and misallocation of talent. The former effect depends on the magnitude of μ_I and μ_{II} which in turn depend on the elasticity of demand in sector I and II..

It is not only the supply responses we should be concerned about but also the

cross-demand effects. Consider two industries where the the products produced by these industries are net substitutes. Then entry barriers in industry I induces, by raising the price of product I, a drop in the quantity of demanded of product I. but also an increase in the demand for the product produced by industry II. This will, depending on the elasticity of supply, increase the price in industry II, making it ceteris paribus more profitable to invest in human capital specific to sector II, however, also in general human capital, as wages increase in both industries. Thus, the supply response, which ceteris paribus lowers wages in industry II, is countered by a demand response that increases the wage, leaving the net effect unclear. Consider now two industries where the products are net complements. Now an increase in the wage and thus the price in industry I decreases demand in industry II, making it less worthy to invest both in human capital specific to industry II, but also in general human capital for those who can enter only industry II. Thus, under net complementarity, the demand response amplifies the supply response's effect on wages, leading to even greater reductions in human capital investment. Suppose now that the the goods produced by the two sectors are neither net substitutes nor net complements. In essence, this could mean that consumers' preferences for the two goods can be represented by a Cobb-Douglas utility function. In this case an increase in the barriers to entry in sector I and the corresponding increase in the price does not induce any demand shift in sector II. The only change will be the positive supply shift in the second sector, which drives down the wage there but the decrease in the wage is not as great as in the case of net complementarity. Thus, the Cobb-Douglas case represents the middle way in assessing the welfare implications of entry barriers. Another important demand response comes from individuals in the protected sector accumulating more human capital which increases their disposable income. Thus they can "offer" more in their trade with the unprotected sector, potentially increasing the income of those in the unprotected sector. intuitively, assigning individuals early to sectors can increase the efficiency of specialization (notwithstanding its negative effect on talent allocation) and thus increase total income.

To put the above expressed intuitions more formally, consider the demand system consisting of two goods produced by sector I and sector II, 1 and 2, characterized with the parameters ϵ_{11}, ϵ_{12}, ϵ_{21}, ϵ_{22}, that is , the own price and the cross price elasticities for both goods. We assume that both goods have an equal share of total spending. Due to symmetry and equal shares, $\epsilon_{12} = \epsilon_{21}$. We will use ϵ_{21} to denote both, as we focus on demand responses in sector II. as a response to price change in sector I. What we will be interested in in particular is the change in the price of good 2 as a result of erecting entry barriers in sector I. As there is both a demand shift and a supply shift, the price change, following basic price theory can be expressed as $\Delta p_2 = \frac{\Delta D_2 - \Delta S_2}{{\epsilon_{22}}^s - {\epsilon_{22}}^d}$. If 1 and 2 are

net complements, it is easy to see that Δp_2 is negative as ΔD_2 is negative, while ΔS_2 is positive. While if they are substitutes the sign of Δp_2 is ambiguous as both ΔD_2 and ΔS_2 is positive. If the former is greater than the latter, Δp_2 will be positive, it will be negative in the contrary situation, while if the two cancel out, it will be zero.

Another important thing to consider is the market size in the protected and in the unprotected sector. Above we have assumed equal shares, but if we drop this assumption, we could interpret the unprotected sector as being an aggregate of many unprotected sectors. The larger this market is, the less will the wage decrease in these sectors as a result of an increase in labor supply, coming from individuals "shut out" of the protected sector. Likewise, demand shifts will also be relatively small. The size of the unprotected sector also matters a great deal for our analysis. Suppose that market demand in sector I is relatively small so that not all members of the privileged class can take a job in it (which is possible if the supply function is flat where it intersects the demand function). Then noone earns a rent in the protective sector which undermines support for the caste system: based on their natural talent and other endowments, some members of the privileged groups could earn more by shifting their labor supply to other sectors better fit for their talents, but cannot do so due to the caste system. If, on the other hand, demand in the protective sector is very large but the caste system is binding, then much of the demand remains unstisfied and thus the deadweight loss from the caste system will be large. This creates pressure against the system, yet there will also be defenders of the caste system: the relatively small number of individuals who, while may not work in the sector fit for their talents, are compensated by a large rent in the protected sector. Now consider an imperfectly enforce caste system (but with an enforced barrier to entry). If demand is high in the protected sector, then, as we have seen, rents are also going to be very high which incentivizes some members of the unpriviledged groups to invest in skills specific to the protected sector. Some of these investments go to "waste" given that most of the unpriviledged will not be able to enter industry I.

As we have already alluded to, there is a tradeoff between two effects of entry barriers to sector I. The first-order effect of the entry barrier is the standard deadweight loss created by the wedge between the demand and the unrestricted supply curve. This deadweight loss, as it is well known is greater the more elastic either demand, supply or both is. However, the second-order effect of the entry barrier, that is, its effect on human capital investment decisions, decreases with elasticity. In sector I a more elastic demand means that the price increases less as a result of the entry barrier, while in industry II, a more elastic demand means the price decreases less with the positive supply response. Also, in industry II, a more elastic supply curve means that as demand increases or decreases the price will not increase or decrease as much as under a more elastic supply

curve.

One implication of our model is that once "discrimination" or exclusion across social groups and markets are present, there will appear differences between groups that can magnify the disadvantage of the disfavored group over time. For instance, blacks or women might be excluded from a profession due to their race or gender, to which they react with investing less in their human capital (due to the supply response in the unprotected sector). This can, over time, create an impression that these groups are somehow naturally inferior workers, and so there will be an added incentive to exclude them, perhaps through statistical discrimination, from the protected sector. These effects are magnified by the intergenerational transmission of human capital (Becker et al., 2018), something we do not model here explicitly. More privileged families will, over time, really have a greater rate of return on human capital investment (both general and specific) as parents' human capital will have a positive effect on their productivity in investing in their children's human capital, so they will invest more, and the excluded (or once-excluded) will invest less. An efficiency-equity tradeoff might develop over time in that redistributing money from the once-privileged to the once-excluded might decrease efficiency as the latter are less productive in investing in human capital. This consideration has important implications for debates about reparations for past injustices: if such reparations are primarily monetary, they may well decrease efficiency (although this is not necessarily the case, see again Becker et al. for further discussion), while investing directly in the human capital of the formerly underprivileged has a lower chance of destroying efficiency and a higher chance of improving it. The above insights can be applied to the case of the accumulation of non-cognitive human capital, that is, the personality traits, habits and values an individual possesses. An underprivileged status leads to less investment in human capital, including non-cognitive skills (which are a special case of general human capital). Thus, historically underprivileged groups will probably have, if parental human capital has an especially strong influence on the production of human capital, a greater prospensity to engage in several forms of "bad" behavior, including crime, harmful addiction and activities generally associated with a lack of self-control. This leads to further deterioration of household human capital production technology. One example could be the propensity to bear children out of wedlock. A large literature exists (see e.g. McLanahan and Sandefur, 1994) on how this negatively affects the production of human capital in children, thus, this helps perpetuate poverty and "underprivileged" status.

Notice that for exclusion to result in differences in income and human capital investment it is not necessary that wages are higher ex ante in the protected sector. While historically one important rationale for erecting barriers to entry based on race, class or gender has been to reserve well-paying jobs to members of the favored group, it is

also possible that the favored group receives higher wages *because* of the entry restriction, which then leads to differences in human capital investments, which then further increase income differences. Arguably this was a case with the exclusion of women from many professions, or the exclusion from blacks in Apartheid South Africa. In fact, one impetus for Apartheid came from trade unions representing white workers (Hutt, 1964), who wished to erect barriers to entry in order to force up their wages. An underdiscussed aspect of this exclusion is that, because whites could foresee that they will enjoy rents in the labor market, they invested more in their human capital, while blacks underinvested in it. In general, we can expect similar results, when job slots in a (perhaps due to union monopoly) restricted labor market job slots are allocated by some arbitrary manner. Obviously, Apartheid has caused a gross misallocation of talent as well as first-order deadweight losses, apart from its effect on human capital investment choices.

Before we turn to another application of our basic model, consider a setup where in both sectors the demand is low enough so that even if only members of a privileged class can enter the industry, they will exhaust all possible rents in the sector. Assume that θ is the probability of group I having exclusive access to the industry, and $1-\theta$ the probability of group II having the same kind of "privilege". Note that in this case it cannot be said to be a real privilege given that no one earns a rent in the industry. The expected wage in the industry for a member of group I can be written as

$$W_I = \theta W_0(H_t, H_I, \underline{H}), \tag{5.14}$$

while in industry II it is

$$W_{II} = (1-\theta) W_0(H_t, H_{II}, \underline{H}). \tag{5.15}$$

An increase in θ ceteris paribus increases specific human capital investment in sector I, and due to complementarity, it also increases general investment in human capital. Misallocation of talent, however, occurs as long as talent distribution in group I and II is overlapping. Note that this case also has perfect equality of opportunity as the first-best efficient case, as in that case everyone simply specializes according to their natural talents. However, once restrictions or social pressures are in place, further tightening (possibly to $\theta = 1$) can be a local optimum. We can imagine a situation where a society starts with relatively small markets but over time demand in individual sectors increases and therefore we move to the case described earlier. Under such a scenario we expect increasing societal pressure directed to abolishing the caste system once markets become larger.

We could further extend our model to consider cases when members of a group have access to multiple industries. Then uncertainty can appear over which sector provides

the best opportunities for a given individuals, which in turn increases the relative share of general human capital investment.

6 Gender norms and the sexual division of labor

An important benefit from family formation throughout history has been the division of labor within the household (Becker, 1981/1991). Typically, women specialized in household work, while men specialized in market work. The pattern of specialization has likely been influenced by comparative advantage (Alesina et al., 2013), however, as Becker (1981/1991, 1985) shows, even without any intrinsic difference between men and women (or a very small difference), specialization emerges in an efficient household. Still, it has always been an open question why specialization within the family has been centered around gender. One argument (made implicitly by Becker, and explicitly by Hadfield (1999)) suggests that gender-based specialization emerged due to "focality". Parents of boys and girls have expected that their child will specialize in household (market) work, so they invested ex ante in household (market) specific human capital. Thus, gender roles have served as solutions to a coordination problem. Another puzzle have been that women have, for a long time, been excluded from certain professions. "Society" enforced such exclusion through both formal laws and social pressure or "conditioning". On the face of it, such exclusion is inefficient as it leads to suboptimal talent allocation. We will show, however, that it need not be the case. More social conditioning means that parents can expect with high certainty that their child will, as an adult, work in the household (market) sector, so they will invest more, and earlier, in household (market) human capital. Perhaps paradoxically, social conditioning can also be efficient when there is little heterogeneity across *individuals* (but there might be heterogeneity across *groups*) in society. We combine these insights in the following short discussion. Our argument resembles Hadfield's argument, with the following important differences: In our model human capital investment is continuous. Also, individuals invest in both sector specific and general skills. Finally, we explicitly discuss the normative (efficiency and equity) implications of our model. Before we turn to the formal analysis let us first consider the case when there are no frictions in the marriage markets and search costs are zero. In such a world, and with the advent of modern dating apps we are increasingly approaching this, there is really no benefit from pre-determined gender roles. For if a man, foreseeing that some women may specialize in market work, accumulates mostly household human capital, he will easily find a woman who is a good match for him. On the other hand, the thickness of the market also matters. A thick market is more likely to offer a wide spectrum of available mates. Thus if search costs are near zero and the

market is sufficiently thick (the letter partly follows from the former), the equilibrium investment in human capital cannot be improved by prescriptive gender roles. Suppose that before an individual enters the marriage market, her parent faces a choice of whether to invest in household or in market specific human capital, apart from human capital that earns a return in both sectors, and the parent also determines the optimal human capital investment. The "wage" in the household sector is determined by innate ability A_h, a "market premium" on household work μ, the share the individual gets from the marital surplus α and the probability that this market premium is realized, θ_h. We assume that θ_h can be influenced by gender norms. For instance, if women are expected to be mothers and homemakers first, then θ_h is relatively large for the parent of a girl and low for the parent of a boy. The wage premium depends on both household-specific and general human capital. Social conditioning might be done through "shaming" the child away from un(wo)manly professions, or, conversely, by employing positive incentives. Obviously, if the former channel dominates, there is an additional welfare cost from conditioning. Here, for the sake of simplicity and brevity, we do not consider such welfare costs. Social conditioning on one side of the market influences expectations on the other side of the market as well. Akin to the case of two-sided markets, conditioning might be applied more forcefully to the side that is more sensitive to conditioning. The expected "wage" in the household sector is

$$W_h = A_h \theta_h \alpha \mu(H_t, H_h, \underline{H}), \tag{6.1}$$

while in the market sector it can be written as

$$W_m = (1 - \theta_h) A_m \alpha W(H_t, H_m, \underline{H}). \tag{6.2}$$

where W is the marginal product of market work.

μ appears due to the specialization decisions on the other side of the marriage market. Thus if men in general invest in market human capital, this increases the benefit of women concentrating on household work. The more men are expected to work in the market and women in the household, the higher θ_h will be for a woman. Note that μ is a "full wage" here, and includes the amount of labor devoted to household work. This is important, as labor-saving technology can decrease the amount of hours devoted to household work, and thus decrease μ (we will return to this later). How does the parent decide how to specialize (on her child's behalf)? She compares the pre-human capital investment expected wage in the household and in the market and chooses to accumulate human capital in the sector with the higher expected pre-investment wage. The parent will invest in household human capital if $A_h \theta_h \mu > (1 - \theta_h) A_m W$, otherwise she invests in

market human capital. Solving for θ_h we get $\theta_h > \frac{A_m W}{A_m W + A_h \mu}$.

After deciding which industry to prepare her child for, the parent then decides on how much money or time to allocate to industry-specific and to industry-neutral human capital investments.

If the parent prepares the child for the household sector, the first order condition with respect to household specific human capital is

$$A_h \alpha \theta_h \mu \frac{dH_h}{dh_h} \left(\frac{\partial W_h}{\partial H_h} + \frac{\partial W_h}{\partial H_t} + \frac{\partial W_h}{\partial H_t} \frac{\partial H_t}{\partial H_h} \right) = \lambda. \tag{6.3}$$

If the parent prepares the child to work in the market, the first order condition with respect to market specific human capital is

$$(1 - \theta_h) A_m \alpha W \frac{dH_m}{dh_m} \left(\frac{\partial W_m}{\partial H_m} + \frac{\partial W_m}{\partial H_t} \frac{\partial W_m}{\partial H_t} \frac{\partial H_t}{\partial H_m} \right) = \lambda. \tag{6.4}$$

The first order condition with respect to general human capital investment is

$$\theta_h A_h \alpha \mu \frac{\partial \underline{H}}{\partial \underline{h}} \left(\frac{\partial W_h}{\partial \underline{H}} + \frac{\partial W_h}{\partial H_t} \right) + (1 - \theta_h) A_m W \frac{\partial \underline{H}}{\partial \underline{h}} \left(\frac{\partial W_m}{\partial \underline{H}} + \frac{\partial W_m}{\partial H_t} \right) = \lambda. \tag{6.5}$$

From this we obtain the "usual" conditions that marginal benefit from specific investment must equal marginal benefit from general investment:

$$\begin{aligned} A_h \theta_h \mu \frac{dH_h}{dh_h} \left(\frac{\partial W_h}{\partial H_h} + \frac{\partial W_h}{\partial H_t} + \frac{\partial W_h}{\partial H_t} \frac{\partial H_t}{\partial H_h} \right) = \\ \theta_h A_h \mu \frac{d\underline{H}}{d\underline{h}} \left(\frac{\partial W_h}{\partial \underline{H}} + \frac{\partial W_h}{\partial H_t} \right) + (1 - \theta_h) A_m W \frac{d\underline{H}}{d\underline{h}} \left(\frac{\partial W_m}{\partial \underline{H}} + \frac{\partial W_m}{\partial H_t} \right) \end{aligned} \tag{6.6}$$

for household investments, and

$$\begin{aligned} (1 - \theta_h) A_m W \frac{dH_m}{dh_m} \left(\frac{\partial W_m}{\partial H_m} + \frac{\partial W_m}{\partial H_t} \frac{\partial W_m}{\partial H_t} \frac{\partial H_t}{\partial H_m} \right) = \\ \theta_h A_h \mu \frac{d\underline{H}}{d\underline{h}} \left(\frac{\partial W_h}{\partial \underline{H}} + \frac{\partial W_h}{\partial H_t} \right) + (1 - \theta_h) A_m W \frac{d\underline{H}}{d\underline{h}} \left(\frac{\partial W_m}{\partial \underline{H}} + \frac{\partial W_m}{\partial H_t} \right) \end{aligned} \tag{6.7}$$

for market investments.

A decrease in uncertainty, that is, a divergence from $\theta_h = \frac{1}{2}$ ceteris paribus, always increases investment in human capital, as long as there is some complementarity between early specific human capital investment and later learning-by-doing, additional to complementarity between general and specific human capital.

Now consider the potential tradeoff between talent allocation and human capital

investment efficiency. Efficient talent allocation would imply that the individual specializes in the market sector if and only if $A_m W > A_h \mu$. However, an increase in the probability of receiving a certain market signal induces more human capital investment as long as there is, on net, complementarity between specific and general human capital, that is, a greater return on specific human capital also induces more (although *relatively* less) investment in general human capital.

Consider a girl who is especially apt in mathematics and not so much in household work, and would like to be a scientist. Clearly, if marriage market considerations would not come into the picture, we could "assign" her a fairly low, even zero, θ_h. However, society cannot assign θ_h individually, only at a societal level, as expectations formed in the other side of the marriage market are encapsulated in θ_h. Therefore, if most girls/women in fact have a comparative advantage in household work, and also when men and women are on average identical and the variance of abilities within genders is small, it is efficient to form strong "gendered expectations", even though some girls or boys will be at the losing end of this. Becker (1981/1991) already showed that even "deviant" girls or boys, that is, children who a gender-nonconformists, can benefit from early gendered socialization, although at a cost, since expectations are formed around the average population values. Our finding here is that it is not only individually rational for parents to do so, but the expectations driving this can be efficient. A change in the average innate abilities of men and women (for example because of the changing nature of market tasks: a move away from the importance of physical strength), however, may change both the efficiency properties and the behavior of households in our "game" as it leads to changing expectations, that is, a change in θ_h. So, for instance, in the case of women's investment in market human capital, an increase in A_m induces more investment in market human capital, which in turn influences the expectations in the other side of the marriage market, leading to a greater $\theta_m = 1 - \theta_h$, which in turn leads to even more investment in market human capital. Presumably, throughout much of human history the gains from the sexual division of labor were so large that society gained a lot from a high θ_h, that is, stronger gendered socialization. In the past few decades this has changed in that, apart from the changing nature of market tasks, there is lower "demand" for stay-at-home mothers' or fathers' time (a low μ), which makes the effect of innate ability larger, as well as increasing its importance.

Now let us look at the role of innate abilities. We first consider the case when there are no innate differences between men and women on average, and also no heterogeneity across individuals. This implies $A_w = A_h$ for any individual. Assume that the individual would get half of the household utility in any role she plays in the marriage. Then specialization then will entirely depend on θ_h. Under $\theta_h = \frac{1}{2}$, there are diminished

incentives to specialize along gender lines, however, if it is either greater or lower than 1, specialization will be early and complete and, with adjusting expectations, θ_h will likely converge to 0 or 1. Thus, gendered expectations always lead to efficient investment in human capital if men and women are on average equally good in market and in household work (as long as individuals simply maximizing their expected income, which we take to be a good simplification when both market and household production is important and human capital intensive) *and* there is little variance across individuals. In such a world gendered expectations are likely to arise spontaneously, as focal points in a coordination game. In fact, Hadfield (1999)'s discussion of gender roles can be regarded as a special case of our model applicable to the case when all individuals are identical. When there are large differences in the average household or market ability of men and women,while there is little heterogeneity *within* genders, and there are high returns from household as well as market specialization, gendered social conditioning is likewise efficient. However, where there is no or little difference on average between the sexes but there is larger within-group heterogeneity, gendered socialization is more costly as it leads to talent misallocation. Why have we seen large changes in gendered expectations in the past few decades? One reason is that it "pays less" to be a household specialist now than it used to be (due to e.g. the emergence of new labor-saving household technology). This, as we have mentioned before, corresponds to a lower μ. A lower μ implies a larger potential role to innate differences and that would imply that more stringent expectations (higher values of θ_h) could reduce expected income as it might lead to human capital malinvestment. Let us now further examine the role of μ. As we mentioned above, μ is the full "wage" one can receive through household work, that is, hours worked in the household are part of μ. How do technological shits effect μ, the propensity to accumulate household human capital as well as the optimal θ_h? Labor-saving household technology, such as washing machines and microwaves induced women to work fewer hours in the household (and more hours in the market). This made it less worthwhile to specialize in household work. A low μ can of course be counteracted with a high θ_h, however, increasing θ_h leads to a lower income than decreasing it (so that parents can react to a dramatic decrease in μ with investing more in market human capital). In fact, if μ is sufficiently low, the optimal θ_h is actually zero, that is, everyone should be expected to specialize in market work. This can be the case when there are fixed costs of investing in household specific human capital. Finally, as we have already alluded to it, a decrease in search costs in the marriage market can also make gender norms obsolete and even dysfunctional as with low search costs even "deviantly" specialized individuals can find a match complementary to their specialization.

One thing we need to consider before we move on with our analysis is that if specific

human capital investment increases, so does the absolute level of general investment. This is because if a parent can be certain that her child will work in the household (market) sector, she is endowed with perfect information that can be used to decide how much specific and general human capital should be accumulated.

7 Comparison with occupational licensing

In many markets occupational licensing laws create a barrier to entry, making the labor supply curve steeper. At first sight such a situation is similar to the cases we emphasized, but there is an important difference. Under occupational licensing who gets part of the rent in the labor market depends on who undertakes the mandated specific human capital investment. It is not known beforehand, unlike in the case of natural talent or under a caste system. However, a similarity is that occupational licensing also increases investment in occupation-specific human capital, only in this case investment is not incentivized, but mandated. Also, it is the requirement of specific human capital investment that makes supply more inelastic in the first place. Once it is inelastic, the incentive to acquire sector-specific human capital does not increase further, as while the marginal return increases, the marginal cost (in terms of foregone earnings) also increases to the same extent. Another difference between the cases we considered and occupational licensing is that in the case of the latter there is a large amount of rent exhaustion. Most welfare analyses of occupation licensing indicate net welfare losses (see e.g. Kleiner and Krueger, 2013, Blair and Chung, 2018), and our analysis does not seem to imply otherwise.

8 Equity-efficiency tradeoffs

Let us now examine more closely whether and when we uncover equity-efficiency tradeoffs in the different cases we discussed. In the case of hereditary or "caste" systems, while there is an equity-efficiency tradeoff in societies where it is very costly to discover and allocate talent using a free market mechanism (a type of society we discussed in subsection 5.1), strictly speaking there is no equity-efficiency tradeoff in a more advanced economy: a level-playing field or "equal opportunity" provides both efficient talent allocation and allows efficient investment in human capital. However, once we have a caste system, a "perfect" caste system can be more efficient than an "imperfect" one. This is because in the former case while there is widespread talent misallocation and total investment in human capital may also be inefficient given that members of the lower castes have diminished incentives to invest, there is an efficient *mix* of human capital investment. In the latter case, however, apart from talent misallocation, there is also substantial human capital

malinvestment.

Let us now proceed to the case of "gendered social conditioning". Here, there is an equity-efficiency tradeoff when either there are large average differences across but not within genders, when there are no or little differences across individuals in general, provided that specialized investments in household and market activity are important enough and there is complementarity between general and specific human capital, as well as complementarity between earlier and later accumulated specific human capital. Apart from incentivizing specialization, gendered social conditioning does not lead to diminished incentives for investment given that increases specialization in one side of the marriage market increases "demand" and incentives for specialization in the other side of the market. The efficiency-equity tradeoff arises as preparing a girl child for household production even though she individually would have a comparative advantage in market production could be construed as unfair. However, when there are large differences in innate ability within genders but not in the average ability across genders, there is little or no equity-efficiency tradeoff, as both equity and efficiency seems to call for a loosening of "gendered expectations".

9 The intergenerational transmission of skills

It used to be common in certain communities that children take up the profession of their parents (or one parent). There are at least two possible, in our view, complementary explanations for that. As we have noted in the introduction, these intergenerational patterns can be the result of a focal point phenomenon: if members of a community expect everyone to follow in the steps of their parents, this can become self-fulfilling. We can also evoke an evolutionary argument: communities where members specialize in this way are (were) more successful than communities where this coordination of plans did not take place and they ended up with say, two bakers, instead of say one baker and one butcher. Another channel though which human capital is intergenerationally correlated is that, as in Becker et al. (2018), parents' human capital influences the technology of investing in children's human capital, creating complementarity between the two. We regard these two channels as complementary: if it is expected beforehand that the son of the baker will himself be a baker, while the son of a butcher will be a butcher as well, the parents will invest in skills accordingly and because they have highly specific human capital, they will be more productive in investing in such human capital in their children. A third channel could be, and this is related to our discussion of quasi-rents, that dynasties are partly the result of parents having "connections" (a type of social capital) in their occupation, which helps their children enter the given industry. Such connections,

or "influence" is likely to be complementary with having human capital specific to the particular sector. Thus the mechanism discussed in Becker et al. and the mechanism described by my essay work in the same direction. The implication of this is that under certain (institutional or technological) conditions, "classes" and dynasties may be even more persistent than in the Becker et al. model.

10 Extension: investments in connections

(Quasi)-rents in labor markets can result from having connections to industry insiders or other forms of social or political influence. Establishing relationships with insiders or accumulating other forms of influence can be considered a form of investment. Such investments may be done either by individuals seeking to establish themselves in a market or by parents who seek to create opportunities for their children. For our purposes the latter variant is the more interesting one. By spending now to create ties with insiders, parents may create future rents for their children. This means, that apart form investing in connection capital, they will also invest more in human capital specific to that industry. The reverse is also true: if parents invest in specific skills, they have greater incentives to invest in connections in the related industry. The two types of human capital thus act as complements.

Investment in "influence" and other human capital may happen at the same time or sequentially. For the sake of simplicity, we will assume that parents first accumulate influence and then invest in the human capital of their offspring. The question arises, which industries will parents accumulate influence? Notice first the complementarity not only between influence and specific human capital, but also between influence and the amount of rent that can be earned through influence.The parent has the strongest incentive to invest in specific human capital for her child In the industry with the most influence. For this reason there is returns to scale in specialization in influence. The parent will optimally choose to invest in influence in only one sector, typically one where the rent in the next period will be the highest. This increases incentives for specific human capital acquisition in that sector, while lowering the share of investments in general skills.

More formally, assume that θ_s is now endogeneous, and is an increasing and smooth function if i, notating the level of influence. Influence can be produced according to a production function $g(i)$, where $g'(i) > 0$ and $g''(i) \leq 0$, assuring an interior optimum. The utility function that the parent maximizes is

$$v = u(C_p) + a\delta \max \left\{ (W_0 + \theta_1(i_1)\mu_1)(H_1, H_t, \underline{H}); ...; (W_0 + \theta_s(i_s)\mu_s)(H_S, H_t, \underline{H}) \right\}. \quad (10.1)$$

As this is a maximum function (parents simply maximize income) and thus can be represented by inverse-L shaped indifference curves, we arrive at a corner solution: the parent chooses, before investing in her child's human capital, the sector with the highest μ_s and will invest in influence only in that sector, implying $\theta = 0$ in every other sector. We can write the first order condition with respect to i as

$$\frac{\partial \theta_s}{\partial i} a \max \delta(W_0 + \mu_s)(H_S, H_t, \underline{H}) = \frac{\partial g(l)}{\partial l}. \tag{10.2}$$

If all μ_s values are the same, the parent chooses randomly,but even then she will choose to invest in connections only in one sector.

We can extend our analysis to consider an equilibrium distribution of influence. Every parent seeks influence in order to benefit their child, but households may have different technologies to invest in influence (some people are better in social networking, or are already part of networks through which they can increase their influence). As more and more households enter the influence contest, the marginal return on influence decreases. At some point, no additional households will enter the "race". Of course, if some of the influence is "hereditary", such as (literally) in the case of hereditary privileges, a number of θ_s values will not equal zero. As in such a case the parent is more likely to invest in the industry with the biggest hereditary θ_s, over generations, influence capital in other sectors will likely depreciate to the point of becoming negligible. However, without inherited influence, and thus only one positive θ_s, the parent will invest in specialization early in the protected sector. As above, it may not mean efficient specialization as long as θ_s is lower than 1: the adult child may, with some probability, end up being shut out of the protected industry and having accumulated the "wrong kind of" human capital.

11 Political economy and societal changes

Our model helps explain a number of societal changes in the past few centuries. As Henry Sumner Maine wrote, the Western world has moved "from status to contract": feudal societies based on status and personal relationships have given way to more meritocratic ways of organizing society. Our model implicates that caste systems are the most unstable when they are not too strict. This is when they lead to the most inefficiencies (both misallocation of talent and misallocation of human capital investments) and this in turn creates pressures for change. Likewise, the changes in the structure of labor demand and household technology have created pressures in more and more occupations to let women enter the labor market. Another prediction that can be drawn from our model is that when "gendered expectations" are present but weak it is especially likely that they will

be overthrown by increasing pressure from interest groups, such as feminist movements. This is because, as we have alluded to it in the relevant section, the expectations might be strong enough to influence talent allocation (in a negative way) while they may not be strong enough to induce a great amount of early investment in human capital. Our analysis has a functionalist flavor as we suggest that both previous laws and norms and the change in them have been "functional" in the sense of contributing to some kind of an efficient organization of productive activities.

12 Concluding remarks

The essay considered the choice problem of individuals (typically parents of children) to invest in human capital that is sector-specific and human capital that is general and increases productivity equally in all sectors. Our findings are quite intuitive: when individuals can reap large returns in a given industry, they will invest more in human capital specific to that industry. Also, if an industry is affected by a positive demand shock at a higher probability, specific investment increases while investment in general human capital may also increase if there is complementarity between the two, although its *share* decreases. The presence of industries where either there exists a large rent on natural talent or where entry is limited leads to more specific human capital investments. Higher uncertainty over future labor market prospects increase the share of general investments and decrease share of specific investments, while potentially leading to less overall human capital investment. We also considered two specific cases: hereditary or "caste" systems and socialization along gender lines. In both cases there can be a tradeoff between investment efficiency and talent-allocation as well as an efficiency-equity tradeoff. In modern economies, this latter tradeoff is less likely to occur, while the former is less stark.

Chapter 4

Supernatural Persuasion in the Family and in Politics

This chapter, co-authored with Barna Bakó, considers the use of religious and other supernatural stories aimed at influencing the behavior of others. In particular, we focus on two contexts where this type of persuasion has been important throughout human history: the family and the relation between a ruler and those rued by him.

1 What can economics teach us about Santa Claus? [1]

1.1 Introduction

In the first part of this chapter we consider a phenomenon so far not treated rigorously by economists: the phenomenon of Santa Claus. Santa Claus presents us a number of intriguing questions: why do children believe in him, why do parents choose to "indoctrinate" children to believe in Santa Claus, why and how is Santa Claus associated with gifts, and more importantly, the withholding of gifts, and finally, how can we make sense of why and when children stop believing in Santa Claus. Are richer or poorer households more likely to experience this phenomenon? What happens when Santa faces competition, from Newborn Jesus Christ, for example? This and related puzzles will be discussed in this section. The first major theoretical results concerning parental influence on children came from Becker (1981/1991) and Becker and Murphy (2000), who describe the socialization of children by parents in economic terms. In their model, a single parent (a mother) is postulated, for the sake of analytical simplicity and it is assumed that she derives utility

[1] This section of the essay was published as 'What can economics teach us about Santa Claus?', *Society and Economy*, 39 (3), pp. 349-358, 2017, with Barna Bak Bakó.

from her own consumption as well as the child's consumption. In the model, the parent can influence her and the child's future consumption and therefore her future utility by instilling certain beliefs and preferences on the child. Becker and Murphy (2000) focus on changing preferences. Their justification for this partial deviation from the standard assumption of stable preferences is the common sense notion that children often do not yet have fully-formed stable preferences and that parents as well as peers influence those preferences and so this is worthy of modeling explicitly. In their model, the parent can induce her children to care for her in her old age, thus one motivation for child-rearing is the support one can get from their children later in life. The parent can do this by investing in persuasion or preference formation that produces guilt in the adult child when he or she does not support his or her old age parent, or producing more altruistic feelings within the child. This result has been controversial as it implies parents sometimes knowingly reduce their child's adult utility in order to increase their own consumption, however, such a result comes easily from any setting where parents are not solely motivated by altruism toward their children. Similarly, the parent can induce other preferences, beliefs, values that either increase her own consumption or the child's consumption later in life. One such investment can be investing in beliefs or preferences for following certain norms that contribute to the adult child's well-being by making the child more well-adjusted and law-abiding in his or her adult life. Part of this indoctrination can be indoctrinating the child with religious or other magical beliefs. Parents have long tried to pass their own religious beliefs and values on to their children, perhaps because they believe it will serve their child well in either their adult life or in an imagined afterlife.

1.2 The basic problem - time inconsistency

Suppose parents wish to incentivize their offspring to do certain things and not do some others. This incentive from the parents' part can come from selfish as well as altruistic considerations: parents may believe (often correctly) that certain types of behavior in the present can improve the outcomes of their children later on, while children are too short-sighted to see the benefits. Alternatively, parents might prefer not being bothered by badly behaving children, so they also might have a "selfish" reason to discipline their children. The distinction does matter when it comes to assessing welfare implications, however, in this article we will focus only on how the parent can incentivize "good" behavior. While the parent cares about incentives, she also considers how it directly affects the child's utility, and the costs of rewards and punishments.

Consider the following full information game parents and their children play with each other. First, the child decides whether to behave well or badly. The parent observes

this behavior and decides whether to give out a reward (R) or a punishment (P) for the observed behavior (assuming that parents will never choose to reward their children for bad, or punish their children for good behavior), or whether to do nothing. Suppose that both reward and punishment involves some costs for the parent. These costs could be the opportunity cost of resources, a time cost of the punishment or the reward, or, in the case of punishment, a psychological cost. Let the cost of reward be C and the cost of punishment be $(1 + a)C$, where a is an exogenous parameter, and- following Becker (1991)'s notation - denotes the extent of parental altruism. Parents receive a payoff B if their child behaves well and a payoff of $-B$ when the child behaves badly (the symmetry of the payoffs are only for reasons of convenience, not necessity). Here we do not consider the problem of whether the parent's utility is derived from selfish or altruistic considerations. We assume that the child has a cost D of good behavior, possibly in the form of unrealized gains from bad behavior and his utility is $-D$ in the case of good behavior if he does not receive any reward for it, and $\delta R - D$ if he receives the reward, where δ is the child's discount factor. His utility is 0 if he behaves badly but does not get punished while he gets $-\delta P$ if he gets punished. Suppose that it is always utility-enhancing for the parent to enforce good behavior, so that $B - C > 0$ and $2B - (1 + a)C > 0$. Moreover, let us assume that $\delta R > D$, so that the reward is sufficient to ensure good behavior.

The unique equilibrium of this game in a one-shot case is that the child behaves badly and the parent does not punish the child. However, we do know that parent-child interactions are typically repeated. In this case the parent can offer the following bargain to the child: I will punish or reward you based on your behavior ever after, and if I fail to live up to my promise, you can punish me by behaving badly. Yet, this repeated parent-child interaction can be best characterized as a finitely repeated game (the parent and the child interact this way as long as the child is dependent on the parent). If the child is rational enough, he knows that in the final round of their repeated interactions the parent has no incentive to reward or punish him, so the cooperation breaks down, not only in the final round but in every period of their interaction.

1.3 A possible solution: creating beliefs

One solution to the above mentioned problem might be to create incentives through persuading individuals of certain supernatural stories. One example might be the widespread cultural phenomenon of the figure of "Santa Claus", named after St Nicholas who is said to distribute presents to children on Christmas Eve (or, in the case of some countries, such as Hungary, on December 6, St Nicholas' Day). It might be the case that parents tell stories about Santa Claus because the consumption of such stories increase the utility of their

children. And therefore, if they are altruistic to some extent, also increase their own utility. However, there are certain features of Santa Claus-type stories that suggest a (perhaps complementary) alternative explanation, the one that is favored by us in this essay: Santa Claus is said to reward "good" children with more and "bad" children with less or no gifts, or perhaps a gift with negative utility, most often coal (or in Hungary, a birch-rod). Santa Claus seems a good candidate for solving parents' time inconsistency problem above: he is implied to have an infinite life, so he has incentives for reputation-building. He is often depicted as someone who cares about social welfare or justice or a similar goal indicating his impartiality. He is said to be relatively unconstrained by the laws of physics and said to be perfectly informed about children's behavior or at least claimed to have very low monitoring costs. More formally, assume for now that children believe in the existence of Santa Claus with probability $\theta = 1$. Put it differently, they believe with certainty that they are playing a game with Santa. Assume that parents tell the following "story" about Santa's utility function: Santa Claus receives a fraction γ of the parent's utility and disutility B and $-B$ and incurs a cost S of distributing gifts where $S < C$ as well as $(1+\mu)S$ for distributing penalties, where μ can be thought of as the extent of Santa Claus' "altruism" toward children. We assume that the decision of Santa Claus is always binary (he either rewards/punishes or does nothing), and the arguments of his utility function are separable. Santa's discount factor is $\delta_s > \delta$. His utility when children in each period behave well and do not get rewarded is $\gamma B + \gamma B/(1-\delta_s)$, $\gamma B + (\gamma B - S)/(1-\delta_s) - S$ when children behave well and get rewarded, $-\gamma B - \gamma B/(1-\delta_s)$ when children behave badly and do not get punished, while it is $-\gamma B + (-\gamma B - (1+\mu)S)/(1-\delta_s) - (1+\mu)S$ when children behave badly and get punished. In a one-shot case, Santa never punishes the child and does not reward the child under good behavior, similarly to the parent. However, suppose Santa's action in each period can be observed by children in the next period. Then if Santa rewards and punishes conditionally, he will get a higher utility in the next period, if and only if $2\delta_s B - (1+\mu)S > 0$. As he plays the game infinitely, the condition can be rewritten as $2B/(1-\delta_s) - (1+\mu)S > 0$. Thus, if Santa Claus is patient enough (a high δ_s) and has low enough reward and punishment cost S and cares relatively little about short-term disutility of children (a low μ), conditional reward and punishment is a best response to good and bad behavior by the child, respectively. To solve for the equilibrium, consider also the child's problem. A given child in any time period has the same payoffs in each case as in the one-shot and finitely repeated game. Since we assumed earlier that $\delta R > D$, a subgame perfect Nash Equilibrium exists in which the child behaves well and Santa Claus rewards the child. The question arises, however, whether the parent has an incentive to play as if she were Santa Claus. In the case when the parent knows for certainty that their interaction ends in the next period, it

is obvious that the logic of finitely repeated games kicks in, that is, the parent will deviate from the "reward when good, punish when bad" strategy in the final round. However, if the parent does not know for certain when the last round occurs, the above described equilibrium can be sustained. The intuition behind this is that if the child has behaved well until a given period but the parent deviates from rewarding the child for it, then the child will modify his beliefs about the existence of Santa Claus and therefore will behave badly with a positive probability after the deviation. As we earlier assumed that incentivizing good behavior is utility-increasing for the parent for any given period, the parent has an incentive to maintain the child's belief in Santa Claus. Note that if this modification of the prior probability occurs, then the game becomes a dynamic stochastic game, with some peculiar features that we will discuss in the following section.

1.4 Uncertainty about Santa Claus and comparative statics

So far we have assumed that children either believe in Santa Claus with certainty or they do not believe in his existence. Naturally, in real life, this is rarely the case among young children. The reason we have not modeled uncertainty is that taking it into account would require us to write down a dynamic and stochastic game where one player (the parent) has an alterego, the child believes with some probability that he plays with the parent or the alterego, which then influences the incentives of the parent. This setup would require a pure theoretic foundation, as the problem itself seems novel in the economic/game theoretic literature. The space of this essay does not allow for such a rigorous treatment, so here we only sketch our intuition as to what comparative statics might be derived from a setup with uncertainty about the existence of Santa Claus. Incentives for supplying beliefs about Santa Claus or similar mythical creatures can depend on the following: if parents are more altruistic, the parent will have lower incentive to punish or not to reward a bad behaving child, so we might expect altruistic parents to spend more resources on creating such beliefs. When alternative means of incentivizing offspring is more costly, we would expect higher reliance on Santa Claus-type stories. We could also expect parents' income to play a role: in poorer families, where children receive less valuable gifts, the prospect of gifts does not constitute such a strong incentive for children than in higher income families. Lower income parents therefore can expect fewer benefits from supplying such stories to their offspring, so we would expect them to rely on other means of disciplining children, such as physical or psychological rewards and punishments more often than higher income parents. Also, if the quality of time spent with children is a normal good, that again reinforces the predicted pattern of richer families relying more on beliefs. We also expect Santa Claus-like stories to be more prevalent in higher income countries. Apart from

the above explanation that is based on the need to reward and punish children, we can propose another reason why parents spread certain beliefs among their offspring. When a parent buys an expensive gift for her child, the child might infer the extent of parental altruism as well as the ability of the parent to buy expensive gifts. Parents might wish to avoid signalling both of these in certain situations. They might prefer their children not to know about their income so that they will not beg for more gifts. Under this account, gift-giving can stem not only from altruism or from the need for providing incentives, but also from pressure from the child. If the child knows that the parent is able and willing to buy him expensive gifts, he has an incentive to pressure the parent to do so (by an embarrassing tantrum, for example). A forward-looking parent might try to avoid such situation by creating a "cover story" for her gift-giving, saying that the gift came not from her, but from, say, Santa Claus. In this account, we would also expect higher income parents to rely more on gift giving, as they have more to lose from pressure-induced gift-giving. Because both explanations predict similar patterns, it is hard to disentangle the effects of the two mechanisms. However, given the evidence of the belief in Santa Claus as an incentive-provider, we believe the first mechanism certainly plays a crucial role, with the later one probably having a complementary role. It is important, however, to stress that traditions involving Santa Claus, like all cultural practices, can have multiple possible "functions". We do have some limited evidence for our intuitions. According to an AP poll performed in the USA in 2006, a higher percentage of children in higher income households believes in Santa Claus than in lower income households. Also, white children are more likely to believe in Santa Claus than black children, though it is unclear how much of the difference can be attributed to income versus other, perhaps cultural factors. The (limited) data so far are consistent with both possible mechanisms described above. Santa Claus is also said to have emerged around the 19th century in Western countries, which again is consistent with our hypothesis that rising incomes increase the demand for Santa Claus related stories through various possible channels. Our comparative statics also predict that as search costs decrease, children will be more likely not to believe in Santa Claus. This explains some of the patterns that we can observe. For instance, many children stop believing in Santa Claus around the time they enter school. One possible explanation for this is that they suddenly acquire more "intellectual maturity". However, to the economist, this explanation is too ad hoc. We find it more likely that as they reach school age, children will interact more with children much older than they are. Older children tend to be better informed for the sheer fact that they have usually spent more time searching for information. If younger children know this, and they also know that fellow children do not have any incentives to make them believe or not believe in Santa Claus, they will interpret their statements as facts increasingly more often than their

parents' words, who, while more informed than even older children, have vested interests in influencing their beliefs. When the incentive disadvantage becomes higher than the information advantage on the margin, children will put a higher weight on the opinion of fellow, older children.

1.5 Conclusions

In this section we have considered the application of basic economic principles to supernatural persuasion within a family context, focusing on the supply and demand for the idea of Santa Claus as a distributor of rewards among children. We have shown how in a one-shot or finitely repeated full information game between a parent and a child the parent has a basic time inconsistency problem when it comes to incentivizing her offspring. This problem can be overcome with the use of supernatural persuasion. Our outline implies that the use of Santa Claus-type stories increases in household income and in children's information costs, and – mainly due to changing information costs – decreases with children's age.

2 Religious persuasion in politics

In this section we consider supernatural persuasion in a different context: we examine the phenomenon of political leaders or rulers using religious persuasion to influence the behavior of their subjects/citizens. We assume that the (sole) leader of a country maximizes net tax revenue, which depends on the tax rate (which is a function of the labor supply decisions of individual citizens), and the wealth or income earned by the citizens. The wealth is a function of some basic institutions, such as the protection of property rights. Protection of property rights is a public good and its production depends on the use of different inputs. One of these are "standard" means of law enforcement, and an other could be religious norms and beliefs, which can stipulate some (divine) punishment as a result of stealing. We assume utility-maximizing agents who make a decision as to how much to steal and how much to engage in productive activities, and in doing so they take into account the benefits as well as the expected costs of stealing vs productive endeavors. We show that, being a public good, there is insufficient incentive to create or perpetuate wealth maximizing religious beliefs, unless the supplier of such beliefs can extract rents from the activity. A political leader with market power has this incentive.

The power of religion in persuading or "indoctrinating" individuals has been recognized throughout the history of humankind. Practically all religious organizations provide instructions to their adherents. These instructions may plant beliefs about the afterlife,

and also contain instructions about how to live. Religion is thus one way through which certain social norms can be internalized by members of a society. These social norms can develop spontaneously but also sometimes deliberately by "norm entrepreneurs". In this chapter we focus on this latter mechanism.

Our essay builds upon a sizable literature in economics. First, we build on the economics of crime, started with Becker's (1968) seminal paper, which aims to analyze criminal behavior as well as punishment as the result of optimizing, rational individual choices. We build on the less well-know literature on the economics of religion, started with Iannaccone (1998). This literature characterizes religion as a market where churches and other religious organizations market their products to consumers, that is, religious individuals. Our essay relates to this literature in furthering the analysis of both the supply and the demand side of supernatural beliefs, concentrating on markets that are monopolistic or at least imperfectly competitive in nature. We also build on the growing literature on the economics and political economy of persuasion. Our approach is especially related to and is building on Becker's and Murphy's (2000) analysis of parental and governmental indoctrination. Becker and Murphy focus on persuasion as a device for changing preferences (or more precisely: changing the marginal benefits for certain activities through persuasion) and thus creating a taste for certain norms. They implicitly build on their earlier work on advertisement as a good or bad. In that model, advertisement and the advertised good are complementary goods, so an increase in each increases the marginal utility from consuming the other. This model also served as an inspiration for us, as we discuss later in more detail. Another approach is what we might call the Bayesian approach. Notable examples include Glaeser's papers on political and parental persuasion, in particular, his work on the political economy of hatred (Glaeser, 2005) or the creation of environmental beliefs and practices (Glaeser, 2014). The problem studied by us is arguably more difficult in that we examine beliefs that are practically impossible to test. In our model we, similar to Becker and Murphy, sidestep the "beliefs vs preferences" issue by simply assuming that persuasion can influence individuals' behavior, either through shaping their preferences and attitudes, or by shaping their beliefs. In the simple version of the model, this distinction is not that interesting, as we shall see. Our essay also has some affinity with the behavioral economics literature, in that the agents in our model seem to depart from those traditionally described by economic models. However, while most of the behavioral economics literature modifies economic models in incorporating certain "biases" or mistakes into individual behavior, which usually mean bounded rationality or imperfect maximization of utility, in our theory individuals are assumed to be utility maximizers a la the standard neoclassical model, however, their beliefs might not correspond to reality (yet, might still be regarded as rational).

Economists have started to systematically study religion only quite recently (Iannacconne, 1998), however, the economics of religion can be said to go back to Smith (1904) and Hume (1985), each of whom stressed the importance of competition and incentives in religious organization and religious "markets". Social science in general has a much longer history of dealing with religion. Apart from explicitly religious thinkers, philosophers have long been fascinated by the possible origins and possible "functions" of religion. Marx (1844) famously dubbed religion an "opium of the people". Durkheim (1912) stressed the role of religion in building communities, or as economists would prefer to call it, producing certain public goods. Weber (1905/1930) has famously theorized about the effects of different religions on economic growth and development, claiming importance for protestant ethic in the development of capitalism. Evolutionary biologists and psychologists stressed the possible adaptive advantages of religion, establishing a view of religion as a natural phenomenon (Dennett, 2006).

Virtually all views on religion seem to agree that it often has a profound effect on human behavior. In economic terms, religion can provide incentives for individuals, and these incentives are often quite important. Belief in heaven and especially belief in hell is correlated with economic growth (Barro and McCleary, 2003). Therefore religion, as it has been suspected by many throughout history, might have an important deterrent effect on certain discouraged behaviors.

Our essay and results are connected to some other traditional problems in the microeconomic literature. Most of all, ours is basically a tradeoff between first-order social welfare losses resulting from a monopoly and the possible welfare gains through the incentives provided by the possibility of market power. This tradeoff has been first, and so far almost exclusively, studied in the context of innovation and intellectual property. Our problem has some similarities with the classic competition/quality innovation problem, only in our case the "innovation" comes from the ability of leaders to produce public goods by using religious persuasion, which they only have sufficient incentives for when they enjoy market power in the political market. In this respect, political market power can be harmful as well as beneficial at the margin.

As we mentioned above, we also build on results in the literature on advertising. Television channels often bundle broadcasted services with advertising, where the viewers value the television program positively, while they may regard advertisements as a "bad". In order to compensate viewers for watching the ads, television companies typically provide their services for free. In our model, it will be churches who act akin to these companies, by providing comfort and other religious goods to believers, together with messages aimed at inducing citizens to engage more in productive and less in unproductive activities.

Throughout much of human history the relationship between religion and the state has been a contentious issue. The church often had large political influence, with the financing of the church often done through mandatory taxes and other contributions. At the same time, the State has had some power over church affairs. The relative influence of Church and State on political and religious affairs has often been the subject of political fights. In our model the church is, unlike in say the Byzantine Empire and like the medieval Catholic Church, nominally independent of the state, however, there is collaboration between church and state. In particular, the ruler "hires" the church to preach messages that are conducive to increasing the tax base, such as "thou shalt not steal!". A particular form of paying the church for partaking in persuasion is to give part of the tax revenue to the church, thus making the church internalize the benefits from religious persuasion.

2.1 The Model

We define the following parameters for our model: a citizen can earn w wage per time spent in the legal sector (x_i), while she can also earn w per time spent in the illegal sector (y_i). An attempted theft is successful with probability η. If the theft is unsuccessful, then the thief is caught and is given a punishment f. A citizen's legally obtained income can be stolen by someone with probability η^{n-1}, where n is the number of citizens.[2] The amount of legal income that is not stolen is subject to a linear income tax τ. Apart from the possible legal punishment, a citizen spending time in the illegal sector also receives a perceived "divine" or "moral" punishment θ, which depends on the level of religious persuasion directed toward the citizenry. We set $\theta = \sqrt{g}$, where g is the amount of resources spent on persuasion. We assume that persuasion occurs through the mediation of the "church". In this we take an approach that is close to how economists model advertising and platform markets. Most television program providers earn their revenues from selling advertisements. The advertisements themselves can either be a "good" or a "bad" from the consumers' perspective (see Becker and Murphy, 1993). If the advertisement is valued negatively by consumers, the platform usually compensates them by offering a service, such as television programs, free of charge. This service is typically "sold" bundled with the advertisement. In our model the leader takes the role of the advertising firm, while the church takes the role of the platform. The church bundles two services: a religious service that is valued positively by churchgoers, and "advertisement" or persuasion that is aimed at the churchgoers as a tool to influence their behavior: in particular, to induce them to allocate more of their time to working instead of predation. Such persuasion

[2]The probability that one particular person steals her income, spending resources y_{-i} on stealing is η, so the probability that *someone* among the $n-1$ other members of society will steal her income is η^{n-1}.

may, for instance, create guilt of a belief that God will punish those who steal either in this life or the afterlife). We do not explicitly model belief formation; instead, we simply assume a persuasion production function.

The church "sells" two goods in a bundle: g and s, where g is persuasion and s is a service provided by the church to its members (such as "peace of mind"). We assume that s is positively valued by churchgoers, while g could influence utility in any direction. In order to keep the analysis simple, and to focus only on the most important aspects, we assume that the intrinsic value of g to churchgoers is exactly zero.

The model's sequence is the following: first, the church decides on how much to spend on religious persuasion and how much on the religious service churchgoers value. Next, the leader chooses the tax rate. Finally, citizens choose how much to work in the legal and in the illegal sector. We first solve for the citizen's optimal choice. The citizen maximizes

$$V_c = w((1-\tau)x_i - \eta^{n-1}y_{-i} + \eta y_i) - ((1-\eta)f + \sqrt{g})y_i - {x_i}^2 - {y_i}^2 \tag{2.1}$$

with respect to x_i and y_i, where x_i is the time and other resources spent in the legal, while y_i is the resources spent in the illegal sector. Both "official" and moral punishment is an increasing function of the time or resources spent on criminal activity. Such a specification is widely accepted in the literature (see e.g. Ehrlich, 1973). Furthermore, the more time and resources one spends on crime, the more resources are "wasted": thus, the social cost of crime increases with the time and resources devoted to it.

The optimal values for each are

$$x_i = \frac{(1-\tau)w}{2} \tag{2.2}$$

and

$$y_i = \frac{\eta(w+f) - f - \sqrt{g}}{2}. \tag{2.3}$$

Note that if η is sufficiently low, y_i would take a negative value, which is "physically impossible". Therefore we make the following restriction: $y_i = 0$, if $\eta \leq \frac{f+sqrtg}{w+f}$.

Taking into account the citizen's future behavior, the leader decides on the applied tax rate. We assume the leader maximizes tax revenue. Tax revenue is the tax rate times the gross domestic product. GDP can be obtained by multiplying x_i by n and w and subtracting from it the amount stolen, which is obtained by multiplying y_i by η w and n. Tax revenue therefore can be written as

$$R = \tau nw\left(\frac{(1-\tau)w}{2} - \eta\left(\frac{\eta(w+f) - f - \sqrt{g}}{2}\right)\right). \tag{2.4}$$

Maximizing with respect to τ we obtain the optimal tax rate

$$\tau = \frac{\eta(f+\sqrt{g}) + w - \eta^2(f+w)}{2w}. \tag{2.5}$$

Taking the choice of the leader into account the church decides on the amount of resources spent on persuasion.

The church gets an α share of the tax revenue. Thus the church's payoff is given as

$$\frac{\alpha nw(\eta(f+\sqrt{g}) + w - \eta^2(f+w))}{4} - c_g g - c_s s. \tag{2.6}$$

The church maximizes the objective function with respect to g and s. The optimal level of g is given as

$$g = \frac{(\alpha\eta nw)^2}{64{c_g}^2}. \tag{2.7}$$

We do not spend time here on the determination of s: s is set simply so as to make it worthwhile for citizens to attend the church services. We can also solve for the optimal tax rate, as well as the optimal x_i and y_i. We obtain

$$\tau = \frac{8w + \eta\left(8f + \sqrt{\frac{(\alpha\eta nw)^2}{{c_g}^2}} - 8\eta(f+w)\right)}{16w}, \tag{2.8}$$

$$x_1 = \frac{1}{32}\left(8w + 8\eta^2(f+w) - \eta\left(8f + \sqrt{\frac{(\alpha\eta nw)^2}{{c_g}^2}}\right)\right) \tag{2.9}$$

and

$$y_1 = \frac{1}{2}\left(f(\eta-1) + \eta w - \frac{1}{8}\sqrt{\frac{(\alpha\eta nw)^2}{{c_g}^2}}\right). \tag{2.10}$$

Finally, we can express the leader's tax revenue, using the above determined equilibrium values, as

$$\begin{aligned}&\frac{8w + \eta\left(8f + \sqrt{\frac{(\alpha\eta nw)^2}{{c_g}^2}} - 8\eta(f+w)\right)}{16w} nw\left(\frac{1}{32}(8w + 8\eta^2(f+w) - \right.\\ &\left.\eta\left(8f + \sqrt{\frac{(\alpha\eta nw)^2}{{c_g}^2}}\right)\right) - \eta\frac{1}{2}\left(f(\eta-1) + \eta w - \frac{1}{8}\sqrt{\frac{(\alpha\eta nw)^2}{{c_g}^2}}\right)\right).\end{aligned} \tag{2.11}$$

We now compute the equilibrium in a case where the leader does not "hire" the

church for persuasion, but instead relies only on physical law enforcement.

The utility function of the citizen is now

$$V_c = w((1-\tau)x_i - \eta^{n-1}y_{-i} + \eta y_i) - (1-\eta)fy_i - {x_i}^2 - {y_i}^2, \tag{2.12}$$

that is, the citizen does not receive a "supernatural" punishment.

The optimal value is the same as in the previous setting,

$$x_i = \frac{(1-\tau)w}{2}, \tag{2.13}$$

while the optimal y_i is

$$y_i = \frac{\eta(w+f)-f}{2}. \tag{2.14}$$

The leader now maximizes

$$R = \tau n w \left(\frac{(1-\tau)w}{2} - \eta \left(\frac{\eta(w+f)-f}{2} \right) \right). \tag{2.15}$$

Maximizing with respect to τ we obtain the optimal tax rate

$$\tau = \frac{\eta(f-\eta(f+w))+w}{2w}. \tag{2.16}$$

Plugging the equilibrium values into the tax revenue we obtain

$$\frac{\eta(f-\eta(f+w))+w}{2w} n w \left(\frac{(1-\tau)w}{2} - \eta \left(\frac{\eta(w+f)-f}{2} \right) \right). \tag{2.17}$$

The church's incentive to invest in persuasion increases in the size of the population, the marginal product of labor, the share the church gets of the tax revenue, and, importantly, on the probability of successful stealing, while it decreases, obviously, in the unit cost of persuasion. Importantly, regarding population size, we find that it is worth for the leader to "rent" the Church's platform if and only if

$$n \geq \frac{8\left(\sqrt{\frac{1}{1-\alpha}}-1\right) c_g (1-\eta)(w+\eta(f+w))}{\alpha \eta^2 w}. \tag{2.18}$$

Essentially there are increasing returns to scale in persuasion. The more people hear the message, the higher will be the benefit from religious indoctrination.

Naturally, we are also interested in the effect of η on whether the leader will rely on the church's persuasive powers. To keep the analysis tractable, we focus on a special case with parameters fixed except for η and n. In particular, we set $c_g = \frac{1}{2}$, $w = 1$, $f = \frac{1}{2}$, and examine the comparative statics for three values of α: $\alpha_1 = 0,1$, $\alpha_2 = \frac{1}{2}$ and $\alpha_3 = 0,9$.

From figure it is clear that the difference between the leader's revenue when he does not rely on the church and his revenue when he does rely on it decreases monotonically in n and η. Intuitively, the harder it is to prevent and prosecute crime, the greater the benefit from relying on persuasion.

Our model presents us a number of further comparative statics results. First, we can say something about the effects of the elasticity of labor supply. The more elastic labor supply is, the lower the equilibrium tax rate, and thus the lower is the leader's incentive to increase the tax base through persuasion. Second, the more difficult it is to enforce laws through physical force (e.g. by maintaining a "police force", corresponding to a higher η), the more does the leader rely on supernatural persuasion. Third, we can also say something about the church's ability to channel supernatural persuasion. The more churchgoers value the free service provided by the church, the more the church will spend on persuasion. Finally, there are some more "obvious" comparative statics implications. Persuasion decreases in the cost of persuasion, and also in the cost of the provision of church service. Although social influences and other ways of upholding public order and incentivizing pro-social behavior are not part of our model, we can hypothesize that when it is harder to observe individuals' behavior, it is more worthwhile to rely on religious persuasion. For a believer, "God sees and knows everything."

2.2 Discussion and applications

Throughout human history we can observe the dynamics and tradeoffs described in our essay. Classic examples of efficient religious norms, such as the "thou shalt not steal" norm highlighted by us have been developed in the ancient Middle East. We know that early civilizations in that territory has typically been based on highly centralized kingdoms with the king having executive, legislative as well as religious functions. Hammurabi's famous laws, often believed to serve as example to many "constitutions" in the ancient Middle East, including parts of the Ten Commandments, can be seen as consistent with our prediction for the case of monopolistic rulers under high barriers to entry. Another corroboration of our main thesis might be the economic history of medieval Europe, a period during which the Catholic Church enjoyed a monopoly and could benefit from economic growth through increased tithe revenues. According to Maddison (2001), economic growth in Europe up until the Black Death have indeed been mildly positive. It is not unreasonable to suggest that without thee Church enforcing efficient norms on property rights, growth would have been even slower or negative.

2.3 Efficiency, Implications for Growth, and Extensions

Is religious persuasion efficient in the Paretian sense? Marx (1844) famously dubbed religion the "opium of the people". Other writers and thinkers also expressed the idea that religion is a tool of the ruling class to oppress the lower classes. This is not the case in our model. In the presented version of the model persuasion benefits everyone as it helps create the public good of "public order", which in turn grows the economy. Pareto-improvement is guaranteed due to the fact that citizens attend church services voluntarily. This feature of our model is similar to the approach of Becker (1996) and Becker and Murphy (2000) who show that even when persuasion is the tool of a ruling class, it is efficient as long as listeners are compensated. In our model the leader might be taken for the ruling class, however, we can easily substitute the leader with a dominant interest group. We can also extend our analysis to a case with a heterogeneous population,like in Becker and Murphy's model. For instance, the rich and the middle class may benefit more from secure property rights than the poor. The poor might even be harmed by them. In this case, the poor would need to be compensated for subjecting themselves to persuasion. Alternatively, in a Bayesian framework, the poor would search more for alternative teachings: they would *like* to believe that stealing is not a sin. In any way, the main results of this chapter would be very similar in such extensions.

Our model seems to suggest that a close connection between Church and State, or a "theocracy" might promote economic development by enabling rulers to rely on religious persuasion. This, however, may be true only in specific cases. If we consider a ruler whose position is "non-contestable", as is the case in our model. Such a ruler clearly has an incentive to promote economic growth as growth increases his tax base. When the political market is contestable, however, leaders may have an incentive to use religious persuasion for different ends. One use of religion could be to differentiate the ruler from challengers. This could create an incumbent advantage in the political market, potentially leading to higher taxes, which then would increase the incentive of the leader to spread efficiency-enhancing messages as well. Computing the equilibrium in such a setting is beyond the scope of this essay, and will likely be the subject of further study. Another important issue worthy of further consideration is the role of the State and the Church in developing human capital. In particular, when human capital is an important driver of economic growth, as is the case especially in modern economies, the leader can have an increased incentive to educate his subjects either directly or through the Church. Investing in human capital can, however, backfire for the leader in the long run, as educated citizens might be more receptive to democratic and other "dangerous" ideas. Having the power to indoctrinate, however, can increase the leader's capacity to invest in education while at the same time preserving his power. In essence, by indoctrinating the

citizens to be loyal to him, the leader can commit the citizens, and this can lead to greater investment in human capital by the leader. Indoctrination can act as a sort of "vaccine" against subversive ideas. Note that as citizens get more educated, it might get harder to indoctrinate them. So the leader faces a fairly complicated optimization problem where spending on the education of the citizenry increases economic output, while potentially leading to democratic change. Spending on indoctrination can help counteracting the latter effect. If the leader's persuasion technology is sufficiently effective, the capacity to persuade or indoctrinate leads to greater per capital GDP. The increase in GDP comes from two effects: the indoctrination of the citizenry and hence providing a "safe route" for investing in human capital, and also through increasing the tax rate, which then provides a greater incentive for the leader to invest in human capital.

2.4 Conclusions

In this chapter we considered the use of religious and other supernatural persuasion in two settings: the persuasion supplied by parents in order to influence the behavior of their children, and persuasion supplied by a leader of a country who "co-opts" the church into using religious persuasion to steer citizens away from unproductive and toward productive activities. In the first context, supernatural persuasion is used in order to solve a commitment problem. In the second setting, it is used especially when laws against theft and other rent-seeking activities are hard to enforce. In both settings, one attractive feature of supernatural persuasion is that supernatural stories usually posit an omniscient being, therefore they can be especially useful when monitoring costs are high.

Chapter 5

Toward an Economics of Moral Character

1 Introduction[1]

In this chapter we propose the integration of ancient theories about moral character (present in what is called virtue ethics as well as in common sense morality) into the human capital literature in general, and the literature on non-cognitive skills in particular.

In the past roughly half a century, the characteristics (skills, knowledge, traits, habits and values) of individuals have come to play a much larger role in economics, due to the emergence of the literature on human capital. In essence, human capital analysis is about how individuals invest in themselves and in their children in order to change their characteristics. Much of the human capital literature has focused on labor market outcomes, such as earnings, but as Becker (1964) have stressed, human capital is also valuable in household production, including leisure. There has also been increases recognition of the fact, than non-cognitive skills also influence individual outcomes and they can also be analyzed as human capital (see the seminal work of e.g. Cunha and Heckman, 2007). Likewise, individuals invest in their health in order to gain life years, better productivity and better quality leisure time. In this essay we build on different strains of the human capital literature, such as the notion of rational addiction (Becker and Murphy, 1988), and the related notion of learning by doing and consumption capital (see also Becker, 1996).

In ancient as well as contemporary virtue ethics and also in much of everyday common sense morality having a good character is often presented as an end in itself. As economists we may be suspicious of such a moral philosophy. After all, if a good moral

[1]The author thanks Barna Bakó and Kevin Murphy for discussions, comments and suggestions.

character leads to good actions, surely it matters that it does so. In the language of economics, the stock of moral character or virtue increases the marginal utility of doing "good" and decreases the marginal utility from doing "bad". Virtue can be acquired by explicit learning or by learning by doing. In either case there is an investment in what we might call "virtue capital". Virtue ethics remain an influential branch of moral philosophy, but it has lost some of the significance it has once enjoyed within philosophy. In fact, virtue ethics was arguably the dominant strain of moral philosophy for centuries or even millenia, up until the birth of modern utilitarianism and deontology.

In general, societies have traditionally placed great emphasis on character traits: good character was to be developed by good parents as well as religious and other social institution. Yet, economists often stress that in dealing with externalities and other social problems, it is the "output" (such as pollution, or crime) that should be "priced" and not the inputs. That is, character might be highly important in so far it leads to better behavior, it is primarily the behavior that should be, in some way, "regulated" by society. If, however, monitoring behavior is sufficiently costly, it may be worthwhile to directly incentivize and monitor the development of personal character. Valued character traits may be rewarded by more success with markets (including in marriage and other social markets) or by praise or shame. To get the reward for one's character traits it is sometimes necessary to "signal" one's character. There is in fact much discussion lately about "virtue signaling". We will see, however, that what is commonly called virtue signaling might signal something other than "virtue". Throughout history, different concepts of "good moral character" have been proposed. In the *Republic*, Plato speaks of the four virtues that are later to be called the "cardinal" virtues: fortitude (courage or bravery), prudence, temperance and justice. Plato thought that temperance was primarily found in the producing class, prudence among the rulers (who need to "manage" things), while fortitude is particularly needed among the warrior classes. The four cardinal virtues also had an important role in medieval Christianity. Pope Gregory VII. also provided a "list" of important virtues. he identified seven virtues that help resisting the temptation of the "seven deadly sins". Clearly, what virtues are productive and for whom has changed over the course of history. Deirdre McCloskey (2006) documents changes in what virtues were commended by society. In particular, during the 17th century, aristocratic virtues has given way to "bourgeois virtues" such as temperance, thrift and honesty.

Both the classical and the early neoclassical economists (such as Marshall (see e.g. Caldari, 2004)) has been much preoccupied with the notion of moral progress. However, as economics has become more specialized, such concerns have been "banished" from economics. Economists have mostly treated individuals and their preferences as given. However, the theory of human capital provides us with a toolkit to study both changes

in people's productive capacities and changes in their preferences or their non-cognitive (personality) traits. Becker (1996) and Becker and Murphy (1988, 2000) successfully applied the human capital framework to study "changing preferences" as well as the influence of the social environment on individuals' choices. In this chapter we essentially take the same approach.. Essentially, following Becker (1996) and Becker and Murphy (1988, 2000), we apply the theory of complements to moral choices.

Our essay complements studies by e.g. Cunha and Heckman (2007), who build and empirically test models of cognitive and non-cognitive skill formation in early childhood and beyond. We are especially indebted to Heckman and others stressing the importance of non-cognitive human capital. A number of non-cognitive skills have been shown to be important in employment, earnings, school behavior, health and other aspects of life outcomes. In this essay we focus more on the non-monetary aspects of non-cognitive traits. We concentrate especially on non-cognitive traits that traditionally has been dubbed as "virtues", and while the literature on non-cognitive human capital concentrates on predicting the social rate of returns as well as the production functions of these skills, we focus on individual maximization. Thus, our approach naturally complements other approaches in this literature. Our focus on the virtues also owes to the observation that individuals often care not only about the actions of others but also their character traits. Within the literature on non-cognitive skills, much work has been done on the effects of "big five" personality traits, such as conscientiousness, agreeableness, openness to experience, extroversion and neuroticism. Conscientiousness, agreeableness, extroversion and emotional stability are usually linked to better outcomes, including in later school performance, adult earnings or (lack of) criminal and other disruptive behavior. One of the chief insight of the economics of non-cognitive traits is that these traits are not fixed at the beginning of life, rather, they can be thought of as forms of human capital: investments can be made in order to build "better personalities" and other desirable character traits and non-cognitive skills. This essay does not have much to add to the literature on the importance of personality traits in general, however, there is much to be done to examine the incentives to invest in such traits in particular settings. We will emphasize traits that has traditionally been described as major "virtues" an individual can possess. We will also consider when societies focus more on outcomes versus character traits, and how the two can, in fact, work as complements. Note that when we talk of investing in non-cognitive human capital, partly we talk about preference formation. One result relevant to our research is Becker and Mulligan (1993)'s work on the endogeneous formation of time preference. In their model, individuals can invest in "imagination capital" which makes them less present-oriented. Higher expected future wages provide incentives for such investment. In our analysis investments in patience will play a role both in individuals'

consumption decisions and their decisions concerning criminal activity. Greater patience demonstrably reduce criminal and other "reckless" behavior as well as engagement in harmful addictions.

The modern work on non-cognitive human capital helps to build a bridge between traditional personality psychology and different strains of virtue ethics. Personality psychologists have long thought that personality cannot be altered, but is "hardwired" into individuals at birth. Philosophers and other earlier thinkers on the other hand have always thought of virtues, that is, traits with some personal or social value, as things to be developed in an individual. Individuals can invest in their own as well as in their children's character traits. What the classical authors lacked was precise knowledge about the mechanisms involved in acquiring character traits. Modern results suggest that character traits become relatively fixed well into the twenties of individuals. Investments in these traits are thus productive mainly in childhood up to early adulthood. There is a clear link between what is called virtue in the philosophical tradition and the personality traits studied by modern psychologists. Take conscientiousness for instance. It clearly has a connection to a number of "virtues" emphasized by both classical and modern authors. One is temperance and prudence: a more conscientious person is better able to withstand the "temptations" of the moment and also to make prudent decisions in everyday life. Emotional stability generally corresponds to prudence, temperance and even fortitude. Honesty is a virtue that is enhanced by conscientiousness.

The essay also relates to the growing literature on preference formation within the family. What we might call moral character often works through preferences, so it is interesting to revisit what we know about this topic. The first forays into this question as been by Becker (1993) as well as Becker and Murphy (2000). Notably, Becker showed that parents, especially less altruistic ones, have incentives to create "guilt" in their child so that the child will be compelled to support them in old age. Interestingly, this actually could be Pareto-improving as it increases human capital investments by parents. Guilt creation is a commitment tool it helps to commit the child to "paying back" the parent for investing in her in her childhood. Following in the footsteps of Becker and others Bisin and Verdier (2001) consider the transmission of cultural traits and the evolution of preferences across generations. Importantly, their model assumes a form of "myopic" altruism were the parent will be "biased" to mold the values and other cultural traits of the child in a direction that resembles the parent's own values and preferences. In both Becker and Bisin and Verdier parents can face a free-riding incentive as children can also learn values from their peers, whose values are influenced by their parents. So parents can, to some extent, free-ride in each other's investments.

One interpretation of this chapter's undertaking is that it is endogenizing individual

preferences. Yet, as e.g. Mulligan (1997) argues, another interpretation can be that investing in certain "values" is complementary to certain behavior. Essentially we can take a household production (Becker, 1965) approach, in which being more "virtuous" increases productivity in "virtuous" acts, while decreases productivity in unvirtuous behavior. The key element in our approach is that whether or not we are talking about changes in preferences, such changes are not sudden and exogenous, but rather they are the results of investment decisions done by parents in their children. Still, individuals choosing e.g. their preferences related to work or drug use might run into problems, as we cannot always easily construct a meta-utility function based on which desirable values are chosen. This presents a general problem with modeling investments in non-cognitive skills: individuals might choose to be more conscientious, but based on what meta-preferences do they decide how much to invest in that trait? We take this problem seriously, and develop ways to solve it. First, we assume that it is parents who invest in their children's trait. They may do so partly due to paternalism, but also because non-cognitive skills increase the productivity of investing in cognitive human capital. But parents can also have selfish reasons to shape their children's attitudes.Becker (1993) presents an example of such parental conduct. We extend his approach in showing that parents may alter their offspring's attitude toward drugs or crime partly in order to decrease their transfers to "broke" adult children. By increasing the hours worked or the wage earned by the adult child, investing in virtues also increases the return to human capital.

Virtue capital investment can be undertaken in various ways. One way in which parents can influence their children's conduct in the future is by setting an example. A parent may, for instance, enjoy smoking but nevertheless refrains fro smoking in order to set a good example to her child. A parent can also simply teach the child or tell stories that have the "right" moral message. It has been established that reading regularly to the child contributes to cognitive as well as non-cognitive development. Virtue can also be accumulated through "learning by doing": doing virtuous acts makes one more virtuous by the force of habit.

The structure of this chapter is as follows. First, we build some intuition for a general framework and comparative statics of investing in moral character. After that, we will present short and simple models that deal with specific applications. In one model we consider parent investing in their child's moral or virtue capital in order to influence the child's future consumption choices in the context of harmful addictions. Then, we consider investing in character that is complementary to law-abiding behavior in the context of crime and punishment. Finally, we build on earlier literature to consider moral character as a commitment device, notably in the context of the organization of work and agency problems. We also consider the issue of "guilt" vs "shame" cultures, as well as the

emergence and decline of "honor cultures".

2 The private and social benefits of moral character

Assume for now that "virtue" is a homogeneous form of human capital. While according to the major ethical systems there are multiple virtues, we concentrate here on some "composite" of these virtue, given that often many virtues contribute to the same outcome and also for reasons of analytic simplicity. Virtue can have many private benefits. First, it can make the individual more patient, which increases her savings and therefore her lifetime income (and indirectly, her lifetime utility), while making her less prone to engage in harmfully addictive behavior. Virtue can also manifest itself as feeling "guilt" when consuming harmful substances or when engaging in behavior that is harmful to others. The latter is, of course, may not be a private benefit, but when social costs are internalized, parents will have an incentive to influence their children's preferences in a way that makes the adult child less prone to engage in crime and other activities that are punished by "society". Virtue also increases the marginal utility of engaging in beneficially addictive as well as socially beneficial activities, including religious and civic activities. When social benefits are internalized, parents will invest in virtue also to make their child more willing to engage in such activities. Finally, by making individuals more honest and hardworking, virtue increases productivity at work.

Let us further dwell on the issue of externalities.The presence of externalities imply that subsidies to invest in virtue capital might be desirable. Subsidizing virtue capital accumulation may not be easy though as governments and societies cannot easily monitor the process. This is where moral persuasion on one hand (the propagation of virtues through social pressure), and direct investments in citizen's virtue might make sense. Often, governments instill virtue capital in their citizens through e.g. the schooling system. We should expect governments and societies to invest more in character building when monitoring and punishing individuals' actions is more costly. Consider for instance medieval Europe. Policing tools were scarce and very expensive, so that states had to rely on very harsh punishment to achieve some crime deterrence. However, criminal punishments typically have an "upper bound": they can be increased a great deal but one cannot go much beyond death by torture, for instance. So under certain circumstances it is very costly to maintain public order by ex-post punishment. In such circumstances it might be a relatively cheap alternative to invest in the preferences, beliefs and character of people.

We may assume (as in Becker and Murphy, 2000) that the marginal cost of investing in people's virtue capital consists of the money income that needs to be paid to citizens in

order to compensate them for their utility loss due to the internalization of norms. When virtue capital has sizable private benefits, this cost is much lower.

3 Ancient versus modern virtues

As we have indicated in the introduction, what character traits are valued by society has changed quite a lot over the centuries. In the Antiquity as well as in the Medieval period traits such as courage have been highly prized, given the frequent need to engage in physical warfare. When recruiting individuals to military units or appointing soldiers to important roles, courageous individuals were rightly favored, while traits such as thrift or empathy were of less value. This has changed about the turn of the Industrial Revolution. As the philosophers of the Scottish Enlightenment, in particular Hume and Smith have stressed, traits such as empathy have come to play a larger role in a commercial society. A businessman who strives to satisfy the wants of his consumers has great use of the skill of being able to imagine himself in the other person's "shoes".

Apart from empathy, other "social"skills have also come to play an important role in commercial societies. The ability to work productively in teams, the ability to "network", to persuade and to control one's emotions in the presence of others are all in general more valuable in a market society than under feudalistic or socialist regimes.

4 Some general comparative statics

In this section we go through some of the predictions of the basic theory of human capital and see how they might apply to the case of moral character. Obviously, conceptualizing virtue as human capital is more useful the more we can apply the fundamentals of human capital theory to it. First, consider time. Investment in virtue should increase in the expected lifespan. This might explain why investing in virtue was often confined to relatively well-off individuals in ancient Greece and Rome, as they had a higher expected lifespan. Here we have complementarities between different types of human capital. More education leads, directly as well as indirectly through a higher expected lifetime wage, to greater life expectancy, which then makes it more worthwhile to invest in other types of human capital as well, including moral character.

Investment in virtue, as far as it is goods intensive, should increase in the income of the parent. This is again a source of complementarity between different forms of human capital.

Investment in virtue capital should decrease in time preference (of those doing the investing).

A higher depreciation rate in the stock of virtue both leads to a lower steady-state level of virtue. A lower level of virtue in the case of the parent has both direct and indirect effects on the child's adult virtue capital. First, it directly influences the production of the child's virtue capital. Second, a lower level of virtue makes the parent spend more on her "bad habits", which drives resources away from investing in her child (this applies to all types of the child's human capital).

An increase in the rate of return to moral character increases an individual's incentive to invest in her own character. However, in our model, it is parents who invest in the virtue capital of their children. In this setting, an increase in the rate of return has an ambiguous effect on virtue capital investment. Parents face a substitution effect: investing in virtue becomes relatively more attractive. However, they also face an income or utility effect, a higher rate of return increase the child s utility which also increases the parent's utility, this would make them spend more on their own consumption.

It is also worthwhile to look at the relationship between virtue and other forms of ensuring good moral conduct. Sociologists and psychologists sometimes speak of "guilt cultures" and "shame cultures" (see e.g. Hiebert, 1985). Guilt cultures make their members internalize certain values, which then influence their conduct. In contrast, shame cultures rely on external incentives, namely the social pressure directed to their members. Simple economic intuition would predict that when it is hard to observe conduct but values can be inferred from conduct that is observed, guilt becomes a more attractive way of influencing human behavior. The same can be said when it requires constant effort to generate "shame", that is, mere public deviance from a norm or group average does not automatically generate shame. Consider again Becker (1992)'s discussion of parents investing in their children's preferences. Parents create guilt in their children which leads adult children to care for the aging parent. The parent could, instead, create shame by continuously reminding the adult child of her duty to care for the parent. The adult child could however, evade this by severing ties with the parent. Guilt, on the other hand, is not diminished by severing ties (it might even increase!). Thus in general, the easier it is to evade social pressure, either because monitoring is hard or either because it is relatively easy to severe ties with one's peers, the more attractive relying on virtue becomes.

5 Character in consumption

There is little doubt, empirically, that parents spend a lot of resources to shape their children's attitudes toward both harmful and beneficial addictions. Parents typically caution against taking drugs, or they might push their children to go to church, to exercise regularly, etc. Yet, it is not entirely clear why altruistic parents would do this. In a Becker

and Murphy (1988) rational addiction framework, individuals become addicts only if it benefits them. If parents are non-paternalistic, that is, they simply take into account their children's utility as the child experiences it, then parents should trust the adult child to make her own decisions. There are four possible reasons for parents to still influence their children's adult choices: first, individuals might be "irrational". This is obviously possible, but here we will focus on explanations consistent with rational choice theory. The second reason is paternalism: while children discount their future utilities, parents may not apply a discount rate across the child's life periods. Thus, parents might care more about the future costs of taking drugs on the well-being of the adult child than the adult child does. "Self-discipline" of an adult child is thus in this model a "merit good", from the parent's perspective. This, however, seems an overly ad hoc assumption. Hence, we introduce a third possibility: harmful addictions lower the income (both the monetary and psychic, and hence, the full income) of the adult child. This induces the parent to transfer resources to the adult/middle aged child, which decreases their own consumption. The parent may avoid this by inoculating the child against consumption of harmful substances and nudging him perhaps toward beneficial addictions. Related to this is the fact that if the per human capital wage of the adult child is lower due to harmful addiction, the return on investment in human capital is lower. The parent can thus increase the return on human capital, and through that her own (altruistic) utility by investing in patience and other virtues in her child.

First, we consider the case of a paternalist parent. We start from and extend the rational addiction model of Becker and Murphy (1988). We analyze three periods. In period one the parent invests in virtue or moral character for her only child, and in the second period, devided into t "subperiods, the adult child makes consumption choices. The return on virtue has multiple dimensions: first, we allow for the possibility that virtue effects adult utility directly, either in a positive or negative way (being virtuous might bring with itself a sense of pride but occasionally also a sense of guilt). Second, virtue decreases the marginal utility of consuming "harmful" or "immoral" goods, which in turn depletes the consumption capital resulting from consuming the harmfully addictive good, and through this, increases adult utility (indirectly). Third, as consumption of harmfully addictive goods can also lower earnings, virtue can have the additional benefit of indirectly increasing the wealth of the adult child.

Assume a utility function with the form $U(x, y, S, V)$, where x is a composite good, y is a good or activity with certain "harmful" properties, S is the stock of past consumption of y, while V is the stock of virtue capital. We establish the following relationships. S affects the marginal utility from consuming or doing y ($U_{yS} > 0$), while the stock of virtue capital decreases its marginal utility ($U_{yV} < 0$). The individual maximizes her utility

subject to the constraint $p_x x + p_y y + p_g g + wt = W(y, V)$, where W is the individual's "full wage" (including both the wage rate as well as hours worked), p_x is the price of the composite good, p_y is the price of the harmfully addictive good, p_g is the "price" of virtue-creation and w is the hourly wage, functioning also as the shadow price of time. Crucially, the wage is also a function of y and V. Consider alcohol or drug abuse. Both will have an affect on the ability to perform work diligently. V also has a direct effect, as people with a higher stock of virtue capital can presumably perform better at work. Virtue may effect total utility positively ($U'(S) > 0$) or negatively ($U'(S) < 0$). For instance, virtue might cause guilt when engaging in "bad" activities. This lowers the marginal utility from "vices" but the guilt could be present even if one does not engage in such activities (see for example the concept of "Catholic guilt" or "Jewish guilt"). Alternatively, being virtuous can feel good, and while one feels guilty when engaging in unvirtuous acts, she feels good otherwise. Notably, even when virtue decreases utility, it can still increase the indirect utility function through its effects on y.

Virtue is produced according to a production function $V(v)$ where v is the amount of time or material resources spent on producing virtue. In the real world, virtue is sometimes accumulated through explicit learning (such as a parent teaching her child about the virtues of honesty and self-control) and sometimes through learning-by-doing. To capture the former we could treat virtue-enhancing activities as being addictive. A good example of such activities is religious practice and the resulting religious capital, which has been studied extensively by Iannacconne (1998) and others. Including learning-by-doing in our model would not, however, lead to different results, hence we focus on a more simple production process.

We solve the model using backward induction, and hence start with the adult child's consumption choices. In doing so we derive the following first-order conditions:

$$\frac{\partial U}{\partial x} = \lambda p_x, \tag{5.1}$$

and

$$\frac{\partial U}{\partial y} + \frac{\partial U}{\partial S}\frac{dS}{dy} = \lambda p_y + \int_0^t e^{-\rho t} p_y. \tag{5.2}$$

where λ is the Lagrange-multiplier and ρ is the discount factor. The first term of the left hand side is positive, while the second term is negative: the explanation is that the consumer gains a positive marginal utility by consuming the harmfully addictive substance, but his current consumption increases the consumption stock S which decreases utility. The right hand side is the present and future marginal cost (price) of the addictive good in each period.

These first-order conditions establish an optimal steady state x and y, x^* and y^*.

The consumption of the numeraire is a function of the lifetime wealth (and indirectly everything else through the Lagrange-multiplier), while the consumption of y is a function of the income, the price of y, the stock of consumption capital S and the stock of virtue V. We assume that in period I. the parent can solve the maximization problem of the future adult child, so she takes the optimal consumptions as given. She maximizes the utility function

$$V_p = U(C_p) + aV_c(x^*(W_c(S)), y^*(W_c(S), p_y, S, V, \beta(V))). \tag{5.3}$$

The first order conditions in the steady state are

$$\frac{dU(C_p)}{dC_p} = \lambda \tag{5.4}$$

and

$$a\frac{dV}{dv}\left(\frac{\partial x^*}{\partial W_c}\frac{\partial W_c}{\partial S}\frac{\partial S}{\partial y}\frac{\partial y}{\partial V} + \frac{\partial V_c}{\partial y^*}\frac{\partial y^*}{\partial V} + \frac{\partial V_c}{\partial S}\frac{dS}{dy}\left(\frac{\partial y}{\partial V} + \frac{\partial y}{\partial \beta}\frac{\partial \beta}{\partial V}\right)\right) = \lambda. \tag{5.5}$$

The first term within the main parenthesis is the increase in consumption due to a greater amount of virtue. The second term is negative, given that $\frac{\partial y^*}{\partial V} < 0$. Marginal utility stemming from the consumption of y decreases as V increases. The third term captures the gain from a lower harmful consumption stock S, and has two parts: the first is a decrease in the consumption stock stemming from the direct effect of V on the consumption of y, while the second one is a decrease due to greater patience. We can conceptualize the effect of V on the consumption of the harmful substance as the parent providing a substitute good for the substance. In the standard rational addiction framework, with a consumption schedule $c(S)$ and a constant depreciation schedule $c = \delta S$, an increase in V, by decreasing consumption at any period t, shifts the consumption schedule downwards, and leads to a lower steady-state consumption level. If the consumption problem has two steady states, and the lower one is an unstable one, the socialization effort could result in the adult child consuming zero units of the harmful substance, or if the child is already a drug addict (moral character building takes place at an older age), this downward shift can cause the (adult) child to stop consuming the harmful substance altogether. This scenario is especially likely in the case of very strong addictions. Hence, the stronger a likely addiction is, the greater is the incentive to influence the adult child's behavior through instilling values or virtues in the child.

Note that the parent maximizes, to the extent of her altruism, the child's lifetime utility without discounting across life periods. This lies behind the parent's incentive to invest in virtuous conduct. For that reason, "good conduct" here is a merit good: if it were entirely up to the (adult) child, he would consume a (perhaps much) greater amount

of the harmful addictive good. Note also how our setup and predictions relating to the discount rate differ from that of Becker and Mulligan. In the Becker-Mulligan model, it is individuals themselves, not their parents, who invests in imagination capital, and so, if they are already addicted to e.g. harmful drugs, it means that their future utility level is lower, while their present utility level is greater than it otherwise would be, and hence, they will invest less in their imagination capital. In our model, it is the parent who makes the choice about investing in patience capital. The parent will take into account the fact that if the adult child becomes addicted to a harmful substance, it will lower her lifetime utility. Increasing the child's patience lowers the chances of the child developing a harmful addiction, and hence, the parent will invest more in imagination capital when the chances of harmful addiction are greater.

5.1 Fully altruistic model with transfers and human capital investments

A merit good or paternalistic approach can at least partially explain why parents invest in "virtues". Parents interested in their child's consumption of merit goods can, as Becker (1981/1991) observes, partially offset the effect of "too much altruism": parents transferring resources to their adult child, which then incentivizes the adult child to be more "careless" and possibly to become addicted to harmful substances, not saving enough money etc. The merit good or paternalism assumptions might, however, seem "ad hoc". In this subsection, instead of postulating that parents do not want their children to become addicted to harmful substances, we derive this "preference" from more basic assumptions. We also introduce the possibility of the parent investing in the child's human capital (mainly: education and health) as well as allowing for the possibility that the parent may transfer resources to her adult child.

The model consists of three periods. In period I the parent raises and socializes the child, and invests in the child's human capital. In period II. the child is a young adult and makes consumption choices, while earning an income. In period II. the adult child is older and "reaps" the (positive and negative) returns of consumption capital.

We assume that the parent is alive throughout the three periods. In period I she is "young", in period II. she is "middle aged", while in period III. she is "old". Analytically, we first solve for the optimal amount of human capital investment and parental transfers in the third period, conditional on the amount of virtue and other variables. Then we solve for the optimal amount of virtue. The marginal benefit of investing in virtue depends on its effect on the marginal return on human capital investment.

Parental transfers

In the last period, the parent obviously does not make any investment decisions, however, she may choose to transfer resources (t) to the middle-aged child. The parent maximizes the lifetime utility function

$$V_p{}^y(C_p{}^y) + V_p{}^m(C_p{}^m) + V_p{}^o(C_p{}^o) + a(V_c{}^y + V_c{}^m), \tag{5.6}$$

subject to the intertemporal budget constraint

$$c_p{}^y + \frac{c_p{}^m}{1+r} + \frac{c_p{}^o}{(1+r)^2} + \frac{t}{(1+r)^2} + h + v = m_p{}^y + \frac{m_p{}^m}{1+r} + \frac{m_p{}^o}{(1+r)^2}. \tag{5.7}$$

We use the intertemporal budget constraint as the parent might transfer the resource in the last period at the expense of consumption of earlier periods. Crucially, we assume that the $V_c{}^m(t)$ fiunction is concave, so the marginal utility of parental transfers (to the parent) is greater when the income of the middle-aged child is lower. This creates an "automatic" incentive for the parent to transfer resources to the child when the child suffers a negative income shock. The first-order conditions yield

$$a\frac{\partial V_c{}^m}{\partial t} = \lambda \tag{5.8}$$

and

$$\frac{dV_p{}^y}{dC_p{}^y} + \frac{dV_p{}^m}{dC_p{}^m} + \frac{dV_p{}^o}{dC_p{}^o} = \lambda, \tag{5.9}$$

which implies

$$a\frac{\partial V_c{}^m}{\partial t} = \frac{dV_p{}^y}{dC_p{}^y} + \frac{dV_p{}^m}{dC_p{}^m} + \frac{dV_p{}^o}{dC_p{}^o}. \tag{5.10}$$

Due to the concavity of the parent's preferences regarding her own and the adult child's consumption, if harmful addiction lowers the child's full income, the marginal utility of transferring resources to the child increases. This, in turn, decreases parental consumption, while at the same time increases the consumption of the harmfully addictive good by the child as he does not bear the full cost of consumption. This problem has at least two possible "solutions": one is that, as in Becker (1981/1991), parental altruism decreases if the child expected to behave ways the parent disapproves of (the merit good case). Alternatively, the parent can spend resources early on to dissuade the child from consuming harmful substances and possibly also to consume goods that are beneficially addictive.

Investment in human capital and in virtue

Apart from possibly transferring resources to the adult child, the parent will also invest in the child's human capital. In period I. the parent decides how much to invest in the child's human capital, while she also decides on how much to invest in his virtue. Resources spent on investing in human capital are notated by h. There is complementarity between the two investments via two channels: first, a greater level of virtue induces the child to avoid harmful addictions which results in a higher lifetime income, which in turn increases the return on (general) human capital. Second, investments in virtue reduces the optimal amount of parental transfers to middle-aged children, which, in turn, increases the optimal amount spent on human capital investment. The first-order condition with respect to human capital investment (in period I.) is

$$a\frac{\partial V_c}{\partial H_c}\frac{dH_c}{dh} = \lambda, \tag{5.11}$$

while the FOC with respect to own consumption is

$$\frac{dV_p{}^y}{dC_p{}^y} + \frac{dV_p{}^m}{dC_p{}^m} + \frac{dV_p{}^o}{dC_p{}^o} = \lambda. \tag{5.12}$$

This implies

$$a\frac{\partial V_c}{\partial H_c}\frac{dH_c}{dh} = \frac{dV_p{}^y}{dC_p{}^y} + \frac{dV_p{}^m}{dC_p{}^m} + \frac{dV_p{}^o}{dC_p{}^o}. \tag{5.13}$$

The condition establishes an optimal level of human capital which depends on h, which in turn depends on the adult wage, and indirectly on the consumption capital stock. The first-order condition for investing in moral character is

$$\begin{aligned} a\frac{dV}{dv}\left(\beta\frac{\partial x^*}{\partial W_c}\frac{\partial W_c}{\partial S}\frac{dS}{dy}\frac{\partial y}{\partial V} + \frac{\partial V_c}{\partial y^*}\frac{\partial y^*}{\partial V} + \beta\frac{\partial V_c}{\partial S}\frac{dS}{dy}\left(\frac{\partial y}{\partial V} + \beta\frac{\partial y}{\partial \beta}\frac{\partial \beta}{\partial V}\right)\right)+ \\ \frac{dV_p{}^y}{dC_p{}^y} + \frac{dV_p{}^m}{dC_p{}^m} + \frac{\partial C_c{}^m}{\partial t}\frac{\partial t}{\partial V}\left(\frac{dV_p{}^o}{dC_p{}^o}\frac{\partial C_p{}^o}{\partial V_c}\frac{\partial V_c}{\partial t}\frac{\partial t}{\partial V}\right)- \\ a\frac{dV}{dv}\left(\frac{\partial V_c}{\partial t}\frac{\partial t}{\partial V} + \frac{\partial V_c}{\partial H_c}\frac{dH_c}{dh}\frac{\partial h}{\partial^2 V}\right) = \lambda. \end{aligned} \tag{5.14}$$

The parent considers three effects of investing in moral character: spending on V induces the adult child to earn more in adulthood directly as well as through incentivizing human capital investments on the parent's part (this is reflected in the last term of the left hand side). This in turn decreases the amount of money transferred by the parent to the adult hild. This directly increases the consumption of the parent (the term in the penultimate parenthesis) but decreases the utility of the child (the first term within the

last parenthesis).

Notice that in the absence of parental indoctrination there would be a contracting problem between the parent and the child. The child would optimally promise the parent to be diligent and not to be a drug-addict in adulthood in exchange for a greater investment in him, however, the child cannot commit to such a deal given that the parent cannot write a contract with the child. A similar problem has been analyzed by Becker: in his model the parent's return on her human capital investments are guaranteed by indoctrinating the child into supporting the parent in old age. A common feature in our model and Becker's is that parents may invest in the socialization or "indoctrination" of their child fully or partly for selfish reasons, which may harm the adult child, however, the child may still benefit on the whole if parental socialization leads to sufficiently higher level of human capital investment. A careful leader might ask, why does an increase in the marginal return on investing in human capital increase the marginal return of investing in virtue capital? After all, according to the envelope theorem, such complementarity should not necessarily be relevant. The reason it is relevant is that without investment in virtue investment in the child's human capital would be *inefficiently* low. Investing in virtue allows the parent to invest efficiently in human capital. Thus, an increase in the return on human capital investments does increase the return on investing in virtue capital. Note that this implies that families which are especially productive in investing in human capital, for instance due to themselves possessing a higher level of human capital, will invest more in their children's virtue.

6 The Persistence of Poverty

A standard model of intergenerational mobility (Becker and Tomes, 1986) posits that at least one potential reason for the persistence of an "underclass" is credit market imperfections in the market for human capital. If household technologies for investing in human capital are uniform across families the poor have a higher rate of return in equilibrium, a source of inefficiency, as they are unable to borrow against future human capital. However, as Becker et al. (2018) show, poverty persistence can result even if there are no credit market imperfections due to differences in household technology. Namely, parents' human capital is an input in the production of children's human capital. Our explanation is complementary mostly to this latter explanation. If the poor are less virtuous on average, then poor families have a worse household production technology for investing in virtue. Furthermore, a credit market imperfection story does not explain any observed gap in the average character traits of rich and poor individuals, as the rate of return on character, most of the time, does not diminish in income.

Thus the key thing that distinguishes poverty persistence as a result of innate differences in willpower and other traits and acquired differences due to bad initial life prospects is how character traits affect wages. It is quite obvious that they do effect wages but the functional form matters a lot. Take the model of Becker and Mulligan (1997). In their model individuals can accumulate "imagination capital" which increases their patience, leading to a higher discount factor. Here the wage effects the marginal product of investing in imagination capital linearly: the higher the wage, the more the individual will invest to increase her patience. Likewise, if a bad habit causes an increase in sick days, the effect is also proportional to the wage. On the other hand, take a bad habit such as gambling. It leads to the individual to spend a lot of money "in vain". The main adverse effect of gambling is that it "wastes" a lot of money. Presumably, poorer individuals probably have more to lose from such an activity. Presumably, possessing a good moral character decreases the marginal utility of many "harmful" activities, such as drinking, gambling, taking drugs, lying, taking advantage of others and so on. The cost from these activities are sometimes increase and sometimes decrease in the expected wage. If the former effect dominates, we expect higher income individuals to invest more in their moral character, while if the latter effect dominates, we expect the poor to be more "virtuous". Alternatively, virtue capital can increase the market wage itself by making the individual more productive. If the second derivative of the wage function with respect to virtue is negative, that is, the marginal returns in investing in virtue decreases in the wage, we again find that the poor have a higher rate of return on virtue capital than does the rich and the middle-class. This is what Adam Smith has alluded to in the Wealth of Nations when he observed that while the upper classes can afford frivolous living, the poor cannot afford it and therefore the latter steer away more often from activities that lead to self/harm, and in particular, financial ruin.

Consider, however, the old landowning aristocracy. Indulging in vices did not lead to a substantial decrease in their material wealth as their income consisted mostly of land rents. Bad habits did cost money but as long as money has declining marginal utility, the landowning class seems like an ideal group when it comes to getting addicted to harmful behaviors. More precisely, we expect people who have a high level of wealth but this wealth does not consist of income from working or investing (think inheritance or land rents) to behave more "recklessly" when the income elasticity of such activities is positive. When the income elasticity is negative, such as in the case of low-quality drugs or low-quality liquors, we expect the relatively poor to behave the most recklessly. Both the income effect (the bad habit being inferior) and the substitution effect (lower lifetime income means a lower expected damage from indulging in the bad habit) points to this direction. Consequently, we may expect members of the "bourgeois class", whose

income mostly consists on wages and returns on physical capital investments, to invest the most in acquiring virtue or good character. Once the lower classes and the "aristocracy" developed bad habits, it can become self-reinforcing as bad character traits lead to bad economic decisions and it also gets passed on across generations. Hence, a subset of the lower classes will remain poor (and form an "underclass") and a subset of the landed aristocracy become impoverished. Another group that is susceptible to low investment in virtue capital is the non-working poor, especially those on welfare and those who expect their children to rely to a large degree on welfare. Welfare payments increase non-work income, thereby exerting an income effect which in itself induces the adult child to engage more in harmful behavior and it decreases time spent at work, decreasing work-related income, thus making it less costly to engage in "vices". It is an open question, one to which we will return later on, whether the poor face credit constraints when it comes to investing in virtue capital. Arguably they do, however, teaching virtue to children might be more time-intensive than goods-intensive.

Note that investments in virtue may pay off even if having a good moral character directly lowers the utility of the adult child. This is because virtue also lowers the marginal utility from engaging in "harmful' behavior, which in turn decreases the "harmful" human capital stock, and that has a positive effect on total utility. If we conceptualize the problem in terms of a household production process (as in Becker, 1965), good moral character increases the shadow price of harmful behavior. If the demand for harmful behavior is sufficiently elastic, this implies that time spent on harmful behavior decreases. In this case the net effect on utility can well be positive, that is, the household will produce more of the household commodity.

Returning to the question of the persistence of poverty, intuitively, poverty persists if the rate of return on virtue capital is permanently low and/or there is a strong complementarity between parents' and children's virtue capital. Take the former case. If the parent expects the child to be low-income as an adult, perhaps due to labor market discrimination or other "external factor", she will invest less in her child's good habits. This way the child will make choices in his adult life that will further plunge him into poverty. On the other hand, if the life prospects of the child increase, the parent invests more in his virtue, and the child has some chance of breaking out of poverty. This will, however, be difficult when parents' virtue influences the productivity of investing in children's virtue capital. If the rate of return is say moderately high but this complementarity between parents' and children's character is strong enough, the parent will, despite a high rate of return, will not invest that much in the child's virtue, and so poverty will persist into the next generation and possibly beyond that. More precisely, as Becker et al. (2018) show, when parental human capital is strongly correlated with the child's human capital, there can be great

persistence "on top" and "on the bottom". That is, there is substantial regression to the mean in middle class families, however, there may be a high degree of persistence among the "underclass" and the "upper classes". This results are obtained when complementarity between parent's and children's human capital is coupled with credit market imperfections (the fact that families cannot borrow against the future human capital of their children). If we translate all this to the case of non-cognitive skills, instead of general moral progress we may see "moral stratification". Another interesting angle is to look at the role of the introduction of social security systems. When old age pensions are available, parents invest less in manipulating their children's preferences (Becker, 1992) as they rely less on the support of their offspring when they are old. As children transfer fewer resources to their parents, the parents face a lower incentive to invest in their children's human capital. Naturally, this includes investing in moral character. Consider a parent with a "weak" moral character and relatively low level of altruism. Even if their child could expect a higher wage in the labor market than the parent, the parent will not internalize this potential change in the family fortune. More "virtuous" families, however, will still invest a lot in their child's virtue given that they are very productive in this type of investment.

Our treatment of the persistence of poverty can potentially solve an important conundrum. Many believe that the poor are poor because they make a lot of bad choices. Others counter that they make bad choices because they are poor (and those choices are rational under adverse circumstances). Our answer is that both of these statements are true. Consider investing in patience. Greater patience will produce more income as it makes the individual invest more in human and physical capital. However, it is also true that if one has worse life prospects, she will rationally invest less in patience.

7 Crime and punishment

Crime has been studied in an economic framework since the seminal contribution of Becker (1968). According to the standard economic model, individuals commit crime when the expected benefits of crime exceed the expected costs. Emphasis has been placed in this literature on studying the deterrent effect of certain punishments as well as the deterrent effect of increasing the probability of punishment by e.g. sending more police officers to the streets. Common observation holds, however, that individuals are also influenced by things like shame or guilt when committing a crime. Here we interpret the guilt felt by potential criminals when committing a crime as a result of the individual possessing moral character. We show that especially when punishments are harsh or the probability of apprehension is high, parents have strong incentives to invest in virtue which in turn

decreases the proclivity of the adult child to commit crimes. As in section 2, we consider two model versions: in version I. the parent is paternalistic in that she does not apply a positive discount rate over her child's life periods. In version II. the parent is completely altruistic, however, as we show, she still has reasons to make her (adult) child guilty about engaging in criminal activity. Let us consider the first case first. Suppose an (adult) individual decides how much crime (x) she engages in. Her utility function is

$$v(y) + b(x) - \beta p f(x) - m(x), \tag{7.1}$$

where b is the (private) benefit from crime, f is the punishment, p is the probability of apprehension,m is the "moral cost" of engaging in crime an β is the discount factor. It can also be interpreted as guilt. We assume that $b'(x) > 0$, $b''(x) \leq 0$, $f'(x) > 0$, $f''(x) > 0$, $m'(x) > 0$ and $m''(x) > 0$. The first order condition with respect to x can be expressed as

$$b'(x) = pf'(x) + m'(x), \tag{7.2}$$

that is, the marginal benefit of crime equals the marginal cost. The optimal level of crime can thus be written as a function $x^*(p, f, m)$, so the indirect utility can be written in the form

$$v(y) + b(x^*(\beta, p, f, m)) - \beta p f(x^*(p, f, m)) - m(x^*(p, f, m)). \tag{7.3}$$

Now allow β and m to be influenced by prior investment in virtue or moral character. Assume a function $m(V)$, with $m'(V) > 0$ and $m''(V) \leq 0$ and a function $\beta(V)$, with $\beta'(V) > 0$ and $\beta''(V) \leq 0$. Let $g(v)$ be the cost of investing in virtue, with $g'(v) > 0$ and $g''(v) > 0$ and v being the input into the virtue production function. As before, we assume that it is the parent who invest in her child's character, and we notate parental altruism again by a. The parent maximizes

$$v(C_p) + a(v(y) + b(x^*(\beta, p, f, m)) - pf(x^*(\beta, p, f, m)) - m(x^*(\beta, p, f, m), V)). \tag{7.4}$$

As in the case of addiction, the parent does not discount across the child's life periods. The first order condition with respect to v is

$$\frac{dV}{dv}\left(\frac{\partial b}{\partial x^*}\frac{\partial x^*}{\partial m}\frac{\partial m}{\partial V} - p\frac{\partial f}{\partial x^*}\frac{\partial x^*}{\partial m}\frac{\partial m}{\partial V} - \frac{\partial m}{\partial x^*}\frac{\partial x^*}{\partial m}\frac{\partial m}{\partial V} + p\frac{\partial f}{\partial x^*}\frac{\partial x^*}{\partial \beta}\frac{\partial \beta}{\partial v}\right) = \frac{\partial g}{\partial v}. \tag{7.5}$$

The bigger is the potential benefit from crime, the lower the marginal benefit of investing in moral character. Notice also that there is some complementarity between f and m. The smaller m is, the greater x will be, which in turn increases the punishment received for committing crime. This way a larger p or f increases the incentive for

investing in virtue by making it more costly to be a "criminal". The last term, however, indicated that there is also substitution between f and m. A higher f decreases x, so a lower m is needed to achieve the same level of deterrence. The former effect dominates when would be criminals' behavior is more responsive to m than to f while the latter dominates in the reverse case. Now let us examine the other way: does investing in virtue by the government (e.g. through "indoctrination" of children in schools) increases the optimal severity or certainty of punishment? The answer is no, as criminal punishment and "moral costs" serve as substitutes in the would-be criminal's utility function. If we introduce a budget constraint for the government: the sate has to allocate funds between punishment and indoctrination, a decrease in the cost of indoctrination would optimally induce more indoctrination and relying less on punishment. On the other hand, the effect of virtue capital on patience is complementary to the amount of punishment, as it is evident from looking at the would-be criminal"s utility function. Thus, if the government initiates a program such as a program applying cognitive behavioral therapy that increases β for potential criminals, the optimal probability as well as the optimal severity of punishment increases for a higher β implies a higher elasticity with respect to p and f. We could further complicate matters by recognizing that investment in moral character by the government partially crowds out character investment by parents. This comes as a social cost when parents are better in developing moral character than are governments. Thus, at least up to a point, it is optimal to rely on stricter or more certain punishment to induce more investment in virtue and through that to ultimately reduce crime, but there may come a point after which investing directly in the virtue capital especially of disadvantaged children takes on a more prominent role. Empirically, the relationship between punishments and investment in moral character implies that the long-run elasticity with respect to criminal punishment can be substantially greater than the short-run elasticity. This effect might pick up only across generations and thus over decades, which makes it hard to empirically test it. Nevertheless, we believe that we have a strong reason to assume that the relationship holds. Increasing the certainty of severity of punishment now will thus possibly have a really large effect on crime in the far future, although it may come at significant cost in the interim period.

The relationship can be more complicated when punishment takes the form of a prison sentence. Serving a prison term lowers the time horizon across which the returns on moral character (such as labor market and household returns) can be collected and thus increasing prison sentences could actually lead to less investment in good morals, habits and other human capital. On the other hand, the threat of punishment ceteris paribus increases labor supply, which increases the utilization rate of most types of human capital.

Now consider the case without parental paternalism, and with the possibility of

(general) human capital investment by the parent. The adult child in this version of the model chooses how much time to devote to crime and how much to work. His utility function can be written as

$$W_i(t_i) + W_l(h, t_l) - \beta p f(t_i) - \beta m(t_i). \tag{7.6}$$

The adult child maximizes his utility function with respect to t_i and t_l. From this we get the indirect utility function

$$\begin{aligned} V_c = W_i(t_i(\beta, p, f, m)) + W_l(h, t_l(\beta, p, f, m, h)) - \\ \beta(pf(t_i(\beta, p, f, m)) + m(V, t_i(\beta, p, f, m))). \end{aligned} \tag{7.7}$$

The parent then takes this indirect utility function and substitutes it into her own utility function:

$$\begin{aligned} V_p = u(C_p) + a[W_i(t_i(W_l(h), \beta, p, f, m(V))) + W_l(h, t_l(W_l(h), \beta, p, f, m(V), h)) - \\ \beta(pf(t_i(, W_l(h), \beta, p, f, m(V))) + m(v, t_i(W_l(h), \beta, p, f, m(V))))]. \end{aligned} \tag{7.8}$$

The parent here is not paternalistic, so she uses the same discount factor the child uses. The parent maximizes utility with respect to v. A higher level of v increases the adult child's time spent working and decreases the time spent on criminal activity. Taking all this as given, the parent chooses how much resources to devote on building the child's human capital. Importantly, we assume that (general) human capital is useful in legal but not in illegal work activity. Thus, we have $W_i'(h) > 0$ and $W_i''(h) < 0$. The first-order condition with respect to h is

$$\frac{\partial W_l}{\partial h} t_l = \lambda. \tag{7.9}$$

As the resources spent on "virtue" increase t_l, they also increase the left-hand-side of the first-order condition. For the equality to hold, $\frac{\partial W_l}{\partial h}$ needs to decrease, which occurs if h increases. We thus showed that investing more in virtue capital leads to more investment in (labor augmenting) human capital. Notice that the amount of human capital in itself leads to a higher t_l (and corresponding lower t_i) chosen, so in a paternalistic model investing in human capital would be another lever the parents could use to discourage their children from crime.

7.1 Revisiting shame and guilt cultures

As we have mentioned above, "guilt" and "shame" are two alternative ways to exert social influence on individuals. We indicated that when behavior is harder to monitor, guilt

will take preeminence over shame. However, just as punishment by law and investing in guilt are complements, so can shame and guilt be complements. When shame is more strictly enforced, parents might try to preempt the shaming of their children by instilling in them guilt in the case of "deviant" behavior. Parents will have an incentive to invest in guilt when the behavioral elasticity with respect to guilt is greater than with respect to shame and if they are sufficiently "paternalistic" toward the child. More formally, consider an individual with a utility function $u(x, y, G, S)$ where G is the "guilt capital" of the individuals. We assume $u_{Gx} > 0$ and thus $u_{xG} < 0$, that is, the more the individual engages in x, the more guilty she will feel, and thus a higher G decreases the marginal utility of engaging in x. The same relationship can be described for S, which is the social capital or social environment of the individual that produces shaming or other forms of social pressure that decreases the marginal utility of engaging in x. Seemingly, the relationship between social pressure and guilt are that of substitutes. A closer examination, however, yields a more complicated picture. Let x^* be the optimal level of x. x^* depends negatively o both G and S.Then a higher expected S has two effects: the (paternalistic) parent can, via ex ante investment, make the effect of a higher S less "painful" to the adult child by decreasing x through an alternative channel. That is, ${x^*}_{SG} < 0$. However, a higher S also partially substitute for the effect of G. That is, if the child has "better" peers, or likely to have better peers as an adult, the parent need not spend as much in instilling good values in her child. Which effect is likely to dominate depends on the behavioral elasticities. If the child's behavior is expected to be elastic with respect to S, it lowers the marginal return of investing in G: less guilt is needed to incentivize good behavior, and also, the child will not suffer greatly from social pressure given her relatively elastic response to it. In the reverse case, that is, when elasticity with respect to the social environment is relatively low, while that with respect to G is relatively high, the marginal return on spending on guilt is relatively high.

The above intuition needs to be modified so far as the peer pressure directed toward the adult child is the result of character traits implanted in other individuals by their own parent. In such a, plausible, case, parents will face an incentive to free-ride on each other's efforts in socializing their children, as discussed by Becker and Murphy (Becker and Murphy, 2000). An increasing importance of social pressure can thus in fact *weaken* the incentive to invest in moral character: character investment by parent i can be a substitute to similar investments by other parents. A solution to this problem, again suggested by Becker and Murphy, is that socialization is partly done by schools, churches and other social organizations, who will appropriately "price" parents' contribution to the development of the child's moral character. If we take into account the endogenous determination of peer pressure, our intuition that when the adult child is relatively

more sensitive to guilt than to social pressure, parents' incentives to invest in guilt is stronger is strengthened: first, the marginal return of creating guilt is higher. Second, the parent can gain a lot by protecting her child against future social pressure and associated distress. And third, parents have relatively little incentive to free-ride on each other's guilt-investments.

8 Honesty as a commitment tool

Commitment problems permeate the economic literature. They have been studied especially in the context of credible deterrence and time-inconsistency in economic policy-making. Commitment, however, is also important in everyday social interactions. Why, for example, people choose "honest" and "trustworthy" individuals as friends and spouses? One explanation could be that they can only imperfectly monitor the other person's behavior: when she is not observed, the individual has an incentive to behave in a "dishonest" manner. If, however, she ex ante invested in virtue capital, she may behave honestly even when she is not observed. Virtue thus serves as a commitment device in ordinary as well as marriage and friendship markets (arguably, virtue can be observed, in selected cases). If the market is competitive, a person with a higher level of virtue can outbid persons with lower levels of virtue given that she can behave honestly in a business transaction, marriage or other partnership at a lower cost, and thus is able to generate a higher surplus.

Investments in honesty and loyalty have been considered e.g. by Akerlof (1983), Akerlof and Kranton (2010), and by Mulligan (1997). Akerlof's model predicts that more honest individuals will get paid more in the labor market given that they are less likely to embezzle or steal while on the job, and therefore, principals trust them more. However, it is not always true that individuals who are known to be honest will earn more. As Akerlof and Kranton, as well as Mulligan show, developing "loyalty" among the agents can help principals economize on their wage bills.

Evolutionary theorists emphasized that traits such as honesty can develop in the course of biological evolution. It should be clear, however, that if traits are important as commitment tools and traits can be altered at least in childhood, there is an incentive to invest in such traits. Becker's above discussed model (Becker, 1993) considers such an investment in terms of parents developing "guilt" in their children to commit the (adult) child to support the parent in old age. We continue in this vain in this section. In particular, apart from Schelling, we follow the lead of Becker and Murphy (2000) who wrote among the first about the importance of "hardwiring" values such as honesty into preferences as a commitment tool.

First, we we consider a simple monetary transaction, with one unit of a good being

offered for sale. The buyer values the good at V, while it costs the seller C to sell the good. Assume $V > C$. Then there is some p at which the transaction should take place. Suppose, however, that the law is weakly enforced, so that either the seller or the buyer can act dishonestly. Let us consider the buyer's position. If she simply takes the good and does not pay for it, she gains p, the price. Without commitment, however, the seller will not be willing to undergo the transaction, as doing so will see her lose C. One way to establish a commitment is to invest in "guilt capital", that produces a level of guilt G if the individual behaves dishonestly. How much G is necessary to achieve full commitment? Obviously, a level that is sufficient to deter the buyer from stealing the good, that is, $G = p$. Suppose it costs D to produce this level of guilt. If there are n possible transactions over the buyer's lifetime, the buyer gains $n(V - p)$ by investing in guilt capital that produces $G = p$ when she behaves dishonestly. Notice that unlike in the case of crime in general, in the commitment problem state punishments and investments in guilt are substitutes. In particular, if there is already a punishment f attached to stealing or behaving dishonestly, the guilt necessary to deter wrongdoing in not p but $p - f$. When f equals p, that is, punishment is calibrated so as to insure full commitment, the parent will not invest in guilt at all. This, of course, does *not* mean that punishment and investments in honesty are substitutes *in general*, as we already showed this in an earlier action.

Another context where moral character can serve as a commitment tool is the labor market. In particular, what personality psychologists call conscientiousness has at least two dimensions through which it effects labor market outcomes. First, a highly conscientious individual has good work habits, which makes her a more productive worker. Its effect is thus the same as the effect of cognitive human capital on wage: as conscientiousness increases labor productivity, it increases the wage the individual is employed at. Another aspect of conscientiousness is a lower proclivity to shirking or malfeasance. This may not affect (directly) labor productivity, however, it makes the individual more trustworthy and acts as a commitment tool. One benefit from employing a conscientious person may be that the employer is able to pay her a *lower* wage, given that "efficiency wages" are not necessary in order to induce effort or discourage malfeasance. Why, then, do parents try to teach their children to be honest in their work? The answer lies in the competitiveness of the labor market. In a competitive labor market an individual worker faces a horizontal (or near-horizontal) labor demand. That is, if she is less honest than other workers and thus the employer has to pay her an efficiency wage, demand for her services will decline to zero. More formally, consider a very simple principal-agent problem with $V(e)$ being the value for the principle derived from the agent attending to her task, w is the wage or other type of payment paid by the principle to the agent, e is the monetary equivalent of

the (effort) cost of attending to the task diligently and $m(V)$ is the monetary equivalent of the "moral cost" of shirking. The moral cost depends on prior investments in honesty. Let p be the probability of the agent finishing the task if she works diligently, and let q be the probability that she will be successful even if she shirks in her duties. The agent's utility if she works diligently is $pW - e$, while if she shirks it is $qW - m(V)$. The agent works diligently if and only if

$$pW - e \geq qW - m. \tag{8.1}$$

Solving for W^*, the wage necessary to incite honest work we obtain

$$W^* = \frac{e - m}{p - q}. \tag{8.2}$$

The wage, unsurprisingly, decreases in m. Now consider a parent's decision to invest in her child's honesty. We assume that the adult child enters a competitive labor market. That is, we can treat W^* as given from the parent's perspective. Given that e, p and q are exogenously given, the only "moving" variable is m. However, as agents are wage-takers, they will take the m already established in the market as given. Let us call this level of m m^*. If an agent has $m > m^*$, and the value of m is common knowledge, she will not get the job. m is, of course, is taken as given by her, but the same is not true of her parent. The parent has the following problem: if she invests sufficiently in her child's "honesty capital", so that $m \geq m^*$, the child will get a job, otherwise the child will not find employment. Let h^* be the amount of honesty that will get $m = m^*$. Then if $W^*l - e(l) \geq C(h^*)$ the parent will teach her child to be honest, otherwise she makes no such investment.

Now consider a labor market with imperfect competition, with the agent having some degree of market power over the principal. Then, the agent does not take W as given. In particular, $W \neq W^*$, instead, a markup is added on W^*. The agent thus faces a downward-sloping residual demand curve. A greater m will both increase the employment chance of the agent, but at the same time it will decrease her wage. Thus, the marginal benefit from investing in honesty is $W'(h) + l'(h) - e(l)$, where $W'(h) < 0$ and $l'(h) > 0$. How much the adult child will gain by being honest depends on the elasticity of demand for her services. Nevertheless, she will gain on net given that a monopolist always operates at the section of the demand curve where demand is elastic. Therefore, we can assume an elastic demand at the relevant interval, thus making investment in honesty always having a positive marginal benefit. The marginal benefit is, however, always lower than in the case of competition. Thus we can arrive at a perhaps not so surprising, but nevertheless novel conclusion: competition increases incentives to invest in one's or one's children's

honesty. The more elastic the agent's residual demand curve is, the more she gains by being known to be honest. Our discussion here can be contrasted with that of Akerlof (1983). In Akerlof's model, having a reputation to be honest *increases* the agent's wage as the principal will know that she is less likely to steal while on the job. In his model, the agent gains by being honest even if demand is relatively inelastic, as both the wage and the quantities sold increase. Which model is closer to reality? In our view it depends on the context. If stealing is the more pressing issue, than Akerlof's model would be more useful, while if shirking and other forms of misbehavior (see Becker and Stigler, 1974) is more prevalent, our model seems more useful.

9 The rise and fall of "honor cultures"

When it comes to externalities, principle-agent problems or collective action problems, it is not only "perpetrators" whose character may be molded as a commitment tool: sometimes it is the "victim" who commits to actions aimed at deterring potential wrongdoings against her. This may be the best economic explanation for the emergence of so-called "honor cultures". In honor cultures when someone gets injured in some ways, one is often expected to take revenge or settle the debate in often violent ways, such as by a duel. The benefit from investing in one's or one's children's "honor capital" increases when alternative ways to defend oneself against wrongdoing are more costly. Laws against murder, rape and other serious crimes may be hard to enforce, so as to make the threat of retaliation a necessary component of protecting individuals against criminal acts. In particular, the sociologist Robert Nisbett (Nisbett and Cohen, 1996) suggested three important ingredients conducive to the development of an honor culture: a scarcity of resources, the presence of situations in which the benefit of theft and crime outweighs the risks, and a lack of sufficient law-enforcement (such as in geographically remote regions). Honor cultures include culture, especially among the aristocracy, in medieval Europe, and more recently the antebellum Southern United States. Presumably, honor cultures can survive even after they lose their "rationale", as parents are better in producing the type of human capital they themselves possess. For example, Southern honor culture has been thought to be persistent even when individuals could already taken their legal disputes to a reliable court system. One explanation for the emergence of Southern honor culture is that the South has been largely populated by herding, as opposed to farming, settlers. Since animals are mobile and hence easier to steal than crops, the threat of retaliation was often the best way to deter theft. Other authors, however, point out that honor culture seems to be stronger in the lowlands, which have focused on producing crops. Another explanations are thus appeared, one of which is that Southern culture is derived

from the honor culture of the British aristocracy (Friend and Glover, 2004). It does seem like honor cultures are more prominent among aristocrats and aristocratic societies than among the middle classes. Here we develop some intuition on why this might have been the case. While aristocrats have had opportunities to go to the court with their legal disputes, they also often had a comparative advantage in the use of violence, given that some form of military training has often been the mainstay of aristocratic boys' upbringing. Military skills are complementary with "honor", given that honor can be more productively defended if one has sufficient skills in using violence. Individuals from the lower classes could likewise invest more in their honor but for different reasons than members of the aristocracy. Unlike the aristocracy, the poor rarely have any military skills. On the other hand, they often face higher costs of using the court system than the middle class, partly due to a lack of education and other human capital. It is thus, for them as well as for the aristocracy (for entirely different reasons) resorting to an honor code is often cheaper than bringing legal matters to the court.

10 Signaling virtue

We have shown in the previous section that when behavior can be cheaply monitored, it is more efficient to incentivize "good" behavior directly, and that if this is being done, individuals and families will have incentives to invest in good moral character. When monitoring is very costly, however, it pays to reward virtue per se. The question emerges, however, how it is possible to verify virtue if behavior cannot be monitored. One answer to this question is that there are aspects of behavior that can be monitored and aspects that cannot be. Consider someone who wishes to marry someone who is altruistic. She may observe the other individual's behavior in the context of e.g. charitable giving. A problem arises though: non-altruistic individuals may have an incentive to "pass" as altruistic: someone may give money to a charity just in order to be seen as altruistic. A separating equilibrium is assured, as usual, if the signal of virtue is costly enough. There is much discussion lately about "virtue signaling". It refers to behavior aimed at demonstrating that one is "virtuous" (according to some measure). However, it is not at all clear that acts of virtue-signaling are actually costly enough signals. Consider online behavior often dubbed as virtue signaling. Someone may retweet a tweet about "racial justice" in order to demonstrate how serious she is about the issue. Yet, the act of a retweet is hardly a costly enough signal to ensure a separating equilibrium. One solution to this puzzle is that such acts do not really demonstrate virtue; rather, they demonstrate allegiance to a social group or "tribe". When we retweet or like something that may anger the outgroup, we thereby do incur a cost and my act can demonstrate my allegiance to

the ingroup.

Singaling virtue is the most advantageous when conduct can be imperfectly monitored, however, occasionally, it *can* be monitored. If it is extremely costly to observe behavior, then virtue cannot be fruitfully signaled and a pooling equilibrium ensues. Therefore, we can expect the most and the least transparent societies to focus the less on personal character. Very transparent societies can focus instead directly on behavior. However, the least transparent societies cannot observe either behavior or character, and these societies suffer greatly from dishonest and other harmful behavior. Such societies must spend heavily on law-enforcement and rely on very high "efficiency wages" to deter dishonest behavior.

One context when signaling virtue seems particularly important is the area of environmental protection. There is increasing social pressure to incite people into behaving an "environmentally conscious" manner. However, much of the relevant behavior is private and thus can only be imperfectly monitored. "Society" therefore may spend resources to "hardwire" environmentally friendly attitudes to individuals, primarily when they are still young. Also, many people would like to reward others for behaving in a "responsible" way. Not being able to monitor their neighbors' private behavior, they look for credible signals of environmental consciousness. As usual, individuals with less "responsible" attitudes might have an incentive to pass as environmentally friendly. Signals, therefore, need to be costly in order to be credible.

11 Interaction with other forms of human capital

So far we have treated forms of human capital other than virtue (such as schooling or health) as given. We could, however, extend our framework to cases where cognitive skills and non-cognitive skills such as virtues interact. Consider two cases: in the first case, poor families face a credit constraint (they cannot borrow against their children's human capital) but have the same household technology, while in the second case they do not face a credit constraint but have a less productive household technology. Intuitively, in both cases the poor invest less in education. The first case is perhaps the most interesting one in that even though investing in virtue is arguably less goods intensive and more time-intensive, and thus credit constraint should not be an issue, it will nevertheless be an issue because when poor families are credit-constrained, their children will earn less and that could lead parents to invest less in their child's virtue capital as well.

To make this point more clear, consider a human capital investment problem with the parent investing in two forms of human capital: cognitive and non-cognitive (the latter can stand for investing in virtue). Suppose that cognitive human capital is produced using money income, while non-cognitive human capital is produced (in the household) exclusively using time inputs. The shadow price of investing in non-cognitive skills is, as it is usual in these problems, the market wage of the parent. The second crucial assumption we make is that there is (some) complementarity between cognitive and non-cognitive human capital. Much will hinge on the strength of this complementarity. If the two forms of human capital are on net substitutes, then a poor parent, if she faces credit constraints, can substitute investing in cognitive sills for investing in non-cognitive skills, which she can do at a lower cost given her relatively low market wage. In the second case, that is, when households have different household human capital production functions, there will be strong complementarities between general, cognitive human capital and non-cognitive skills, including moral character. Even if a household is only "bad" at investing in one type of human capital, complementarities across types of human capital imply that the household will also invest relatively little even in skills where they do not have an "inferior" technology.

12 Conclusions

In this chapter we analyzed investments by parents in their children's values or "virtue capital". In particular, we sketched three simple models of such investments: one considering investments in good consumption habits, one analyzing the interaction of criminal punishment and the creation of attitudes about crime and one dealing with virtue capital as a

commitment device. We also provided a reconsideration of what are termed "guilt", "shame" and "honor" cultures. In summary, we pointed out a number of complementarities between various ways of influencing human conduct, and between various forms of (cognitive and non-cognitive) human capital. The models we presented can each be improved and extended further, hence, this chapter serves as a starting point for potential future research.

Summary

In each of the preceding five essays we presented models of applying the economic way of thinking with everyday problems, as well as problems not yet analyzed in the economic literature. In the first essay we considered the problem of information aversion in a market context. Information aversion has been shown to be important in a number of areas. The novelty of our essay is the put the problem into a context of market equilibrium. We showed that information aversion induced misinformation can be an equilibrium outcome, however, how strongly it manifests itself is sensitive to market incentives as well as the opportunity to use the court system.

The second essay considered the effects of the exit of Uber from the Hungarian market on bicycle sharing usage. Somewhat counterintuitively, we found that the exit decreased BSS usage among regular users. This suggests that many use bicycle sharing as part of a “multimodal” pattern of transportation use.

The third essay is concerned with investments by parents in general and industry-specific human capital for their children. In particular, we found that how early specialization takes place depends on uncertainty over future potential rents. The main implications of the model is that both very high achievers in one particular area ("superstars”) and general low-achievers will specialize relatively early, while general "good students” will delay specialization and invest relatively more in general skills. Both superstar specialization and general investments increase in market size as well as in better information flows. Our investigations led to us uncovering interesting tradeoffs between investment efficiency and talent allocation in certain cases, and shed light on the history and practice of hereditary occupations and the sexual division of labor.

In the fourth chapter we first examine the incentives of parents to create supernatural beliefs (such as a belief in Santa Claus) in their children. In our model parents do so in order to influence their offspring’s behavior. Supernatural beliefs help parents to overcome what otherwise would be a commitment problem on their part. In the remainder of the chapter we sketch a model of supernatural persuasion by the leader of a country. In the model the leader "co-opts” "the Church” to bundle religious services with messages aimed at discouraging citizens from stealing and other unproductive activities and steer them

toward productive activities. In particular, the harder it is to enforce laws, the more leaders will rely on religious and other forms of persuasion.

The last chapter contains research that is a work in progress. It presents models analyzing parents' decisions to build "moral character" in their children. Investing in these character traits becomes more important when individuals' actions are harder to observe. Furthermore, we show that there exist interesting complementarities between character-building and other ways of influencing individual behavior. For instance, stricter punishments may incentivize parents to steer their child away from becoming a criminal in the future. There is also complementarity between character traits and general human capital. For example, steering the child away from using drugs as an adult increases the lifetime income of the child which in turn increases the rate of return on general human capital investments.

Finally let us mention a few directions in which the research presented in this thesis can be improved upon. We are still working on providing a model of specific and general human capital accumulation that uses explicit functions instead of the implicit ones used in Chapter 3. Such a move would sacrifice some generality, but would add tractability and would yield more precise comparative statics results. Chapter 4 would benefit from extending the analysis of religious persuasion to include the provision of educational services by governments and churches. Finally, the results from Chapter 5 could form the bases of more than one future publications. It is for that reason as well that we welcome any comments, criticisms and suggestions on this draft of our thesis.

Bibliography

Akerlof, G. A. (1983), ' Loyalty Filters', *American Economic Review*, 73 (1), pp. 54- 63

Akerlof, George A. and Dickens, William T. (1982), 'The Economic Consequences of Cognitive Dissonance', *American Economic Review*, 72 (3), pp. 307-319

Akerlof, G. A. and Kranton, R. (2010), *Identity Economics*, Princeton University Press

Alesina, A., Giuliano, P. and Nunn, N. (2013), 'On the Origins of Gender Roles: Women and the Plough', *Quarterly Journal of Economics*, 128 (2), pp. 469-530

Allison, P. D. and Waterman, R. P. (2002), 'Fixed-Effects Negative Binomial Regression Models', *Sociological Methodology*, 32, 247-265.

AP Poll: Santa Claus Endures in America, in The Washington Post, by Calvin Woodward, The Associated Press, Saturday, December 23, 2006

Barro, R.J. and McCleary, R.M. (2003), 'Religion and Economic Growth across Countries', *American Sociological Review*, 68 (5), pp. 760-781

Becker, G. S. (1996), *Accounting for Tastes*, Harvard University Press

Becker, G.S. (1965), 'A Theory of the Allocation of Time, *The Economic Journal*, 75 (29), pp. 493-517

Becker, G. S. (1968), 'Crime and Punishment: An Economic Approach', *Journal of Political Economy*

Becker, G.S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, NBER, New York

Becker, G. S. (1985), ' Human Capital, Effort, and the Sexual Division of Labor', *Journal of Labor Economics*, 3 (1), pp. 33-58

Becker, G.S. (1993), 'Nobel Lecture: The Economic Way of Looking at Behavior', *Journal of Political Economy*, 101 (3)

Becker, G. S. (1981/1991), *Treatise on the Family*, Harvard University Press

Becker, G.S; Kominers, S.D.; Murphy, K. M. and Spenkuch, J.L. (2018), 'A Theory of Intergenerational Mobility', *Journal of Political Economy*, 126 (1)

Becker, G. S. and Mulligan, C. B. (1997), 'The Endogenous Determination of Time Preference', *Quarterly Journal of Economics*, 112 (3), pp. 729-758

Becker, G. S. and Murphy, K. M. (1993), 'A Simple Theory of Advertising as a Good or Bad', *The Quarterly Journal of Economics*, 108(4), pp. 941-964

Becker, G. S. and Murphy, K. M. (1988), 'A Theory of Rational Addiction', *Journal of Political Economy*, 96(4), pp. 675-700

Becker, G. S. and Murphy, K. M. (2000), *Social Economics: Market Behavior in a Social Environment*, Harvard University Press

Becker, G. S. and Murphy, K. M. (1992), 'The Division of Labor, Coordination Costs, and Knowledge', *Quarterly Journal of Economics*, 107 (4), pp. 1137-1160

Becker, G.S. and Stigler, G.J. (1974), 'Law Enforcement, Malfeasance, and Compensation of Enforcers', *Journal of Legal Studies*, 3 (1), pp. 1-18

Becker, G.S. and Tomes, Nigel (1986), 'Human Capital and the Rise and Fall of Families', *Journal of Labor Economics*, 4 (3), pp. 1-39

Bisin, A. and Verdier, T. (2001), 'The Economics of Cultural Transmission and the Dynamics of Preferences', *Journal of Economic Theory*, 97 (2), pp. 298-319

Blair, P. Q. and Chung, B. W. (2018), 'How Much of Barrier to Entry is Occupational Licensing?', NBER Working Paper No. 25262

Caldari, K. (2004), 'Alfred Marshall's Idea of Progress and Sustainable Development', *Journal of the History of Economic Thought*, 26 (4), pp. 519-536

Campbell, K. B. and Brakewood, C. (2017), 'Sharing Riders: How Bikesharing Impacts Bus Ridership in New York City', *Transportation Research Part A*, 100, pp. 264-282

Clauss, T. and Döppe, S (2016), 'Why Do Urban Travelers Select Multimodal Travel Options: A Repertory Grid Analysis', *Transportation Research Part A*, 93(C), pp. 93-116

Cohen, P.; Hahn, R.; Hall, J.; Levitt, S. and Metcalfe, R. (2016), 'Using Big Data to Estimate Consumer Surplus: The Case of Uber', NBER Working Papers 22627

Cowen, T. (2005), 'Self-Deception as the Root of Political Failure', *Public Choice*, 124(3-4), pp. 437-451

Cowen, T.; Glazer, A. and McMillan, H. (1994), 'Rent Seeking Can Promote The Provision of Public Goods', *Economics and Politics*, 6 (2), pp. 131-145

Crainic, T. G. and Kim, K. H. (2007), *Intermodal Transportation*, in Handbooks in Operations Research and Management Science, Vol. 14, pp. 467-537

Cramer, J. and Krueger, A.B. (2016), 'Disruptive Change in the Taxi Business: the Case of Uber', *American Economic Review*, 106(5), pp. 177-182

Cunha, F. and Heckman, J. J. (2007), ''The Technology of Skill Formation', *American Economic Review*, 97 (2), pp. 31-47

Dawson, E., Savitsky, K. and Dunning, D. (2006), '"Don't Tell Me, I Don't Want to Know": Understanding People's Reluctance to Obtain Medical Diagnostic Information', *Journal of Applied Social Psychology*, 36(3), pp. 751-768

de Chardon, C. M.; Caruso, G. and Thomas, I. (2017), 'Bicycle-Sharing System 'Success' Determinants', *Transportation Research Part A*, 100, pp. 202-214

Deming, D. J. and Noray, K. (2020), 'Earnings Dynamics, Changing Job Skills, and STEM Careers', *Quarterly Journal of Economics*, forthcoming

Dennett, D. C. (2006), *Breaking the Spell: Religion as a Natural Phenomenon*, Viking (Penguin)

Dixit, Av. and Norman, V. (1978), 'Advertising and Welfare', *The Bell Journal of Economics*, 9(1), pp. 1-17

Durkheim, E. (1912), *The Elementary Forms of the Religious Life*, Allen and Unwin, London

Ehrlich, Isaac (1973): Participation in Illegitimate Activities: A Theoretical and Empirical Investigation, *Journal of Political Economy*, 81(3), pp. 521-565

El-Assi, W.; Mahmoud, M. S. and Habib, K. N. (2017), 'Effects of Built Environment and Weather on Bike-Sharing Demand: a Station Level Analysis of Commercial Bike-Sharing in Toronto', *Transportation*, 44, pp. 589-613

Faghih-Imani, A. and Eluru, N. (2015), 'Analysing Bicycle-Sharing System User Destination Choice Preferences: Chicago's Divvy System', *Journal of Transport Geography*, 44, pp. 53-64

Faghih-Imani, A. and Eluru, N. (2016), 'Determining the Role of Bicycle-Sharing System Infrastructure Installation Decision on Usage: Case Study of Montreal BIXI System', *Transportation Research Part A*, 94, pp. 685-698

Faghih-Imani, A.; Aanowar, S.; Miller, E. J. and Eluru, N. (2017a): Hail a Cab or Ride a Bike? A Travel Time Comparison of Taxi and Bicycle-Sharing Systems in New York City. *Transportation Research Part A*, Vol. 101, pp. 11-21

Faghih-Imani, A.; Hampshire, R.; Marla, L. and Eluru, N. (2017b), 'An Empirical Analysis of Bike-Sharing Usage and Rebalancing: Evidence from Barcelona and Seville', *Transportation Research Part A*, 97, pp. 177-191

Fernandez-Heredia, A.; Jara-Diaz, S. and Monzon, A. (2016), 'Modelling Bicycle Use Intention: the Role of Perceptions', *Transportation*, 43, pp. 1-23

Fishman, E. (2016), 'Bikeshare: A Review of Recent Literature', *Transport Review*, 36, pp. 92-113

Fishman, E.; Washington, S.; Haworth, N. and Watson, A. (2015), 'Factors Influencing Bike Share Membership: An Analysis of Melbourne and Brisbane', *Transportation Research Part A*, 71, pp. 17-30.

Ganguly, A. and Tasoff, J. (2017), 'Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future', *Management Science*, 63(12), pp. 3999-4446

Gebhart, K. and Noland, R. B. (2014), 'The Impact of Weather Conditions on Bikeshare Trips in Washington, DC', *Transportation*, 41, pp. 1205-1225

Gentzkow, M. and Shapiro, J. M. (2006), 'Media Bias and Reputation', *Journal of Political Economy*, 114(2), pp. 280-316

Gervais, M., Livshits, I. and Meh, C. (2008), 'Uncertainty and the Specificity of Human Capital', *Journal of Economic Theory*, 143 (1), pp. 469-498

Glaeser, E. L. (2005), 'The Political Economy of Hatred', *Quarterly Journal of Economics*

Glaeser, E. L. (2014), 'The Supply of Environmentalism: Psychological Interventions and Economics', *Review of Environmental Economics and Policy*, 08 (02), pp. 208-229

Glaeser, E.L., Johnson, S. and Shleifer, A. (2001), 'Coase versus the Coasians', *Quarterly Journal of Economics*, 116 (3), pp. 853-899

Glaeser, E. L. and Shleifer, A. (2003), 'The Rise of the Regulatory State', *Journal of Economic Literature*, 41(2), pp. 401-425

Glaeser, E. L. and Ujhelyi, G. (2010), 'Regulating misinformation', *Journal of Public Economics* 94(3-4), pp. 247-257

Golman, R., Hagmann, D. and Loewenstein, G. (2017), 'Information Avoidance', *Journal of Economic Literature*, 55(1), pp. 96-135

Gonzalez, F.; Melo-Riquelme, C. and De Grange, L. (2016), 'A Combined Destination and Route Choice Model for a Bicycle-Sharing System', *Transportation*, 43, pp. 407-423

Goodman, A. and Cheshire, J. (2014), 'Inequalities in the London Bicycle-Sharing System Revisited: Impacts of Extending the Scheme to Poorer Areas but then Doubling Prices', *Journal of Transport Geography*, 41, pp. 272-279

Grusky, O., Roberts, K. J. and Swanson, A-N. (2007), 'Failure to Return for HIV Test Results: A Pilot Study of Three Community Testing Sites', *Journal of the International Association of Providers of AIDS Care*, 6(1), pp. 47-55

Hadfield, G. K. (1999), 'A Coordination Model of the Sexual Division of Labor', *Journal of Economic Behavior and Organization*, 40 (2), pp. 125-153

Hall, J. D.; Palsson, C. and Price, J. (2017), 'Is Uber a Substitute or Complement for Public Transit?', Working Paper.

Hampshire, R. C.; Simekb, C.; Fabusuyia, T.; Dic, X. and Chen, X. (2017), '*Measuring the Impact of an Unanticipated Disruption of Uber/Lyft in Austin, TX*', Working Paper.

Hiebert, P. G. (1985), *Anthropological Insights for Missionaries*, Grand Rapids: Baker Book House

Horan, H. (2017), 'Will the Growth of Uber Increase Economic Welfare?', *Transportation Law Journal*, 44, pp. 33-105

Howell, J. L. and Shepperd J. A. (2012a), 'Reducing Information Avoidance Through Affirmation', *Psychological Science*, 23(2), pp. 141-145

Howell, J. L. and Shepperd J. A. (2012b), 'Behavioral Obligation and Information Avoidance', *Annals of Behavioral Medicine*, 45(2), pp. 258-263

HCSO (2018), *Statistical mirror - Health Status and Health Lifestyle Behaviour, 2016–2017*, Budapest, Hungarian Central Statistical Office.

Hume, D. (1985), *Essays, Moral, Political and Literary*, revised edition by E.F. Miller, Indianapolis: Liberty Classics

Hutt, W.H. (1964), *The Economics of the Colour Bar: A Study of the Economic Origins and Consequences of Racial Segregation in South Africa*, Andre Deutsch Ltd

Iannaccone, L. (1998): Introduction to the Economics of Religion, *Journal of Economic Literature*, Vol. 36, pp. 1465-1496

Jaffe, S., Minton, R., Mulligan, C. B. and Murphy, K. M. (2019), *Chicago Price Theory*, Princeton University Press

Johansson, C; Lövenheim, B.; Schantz, P.; Wahlgren, L.; Almströmd, P.; Markstedt, A.; Strömgren, M.; Forsberg, B.; Sommar, J.N. (2017), 'Impacts on Air Pollution and Health by Changing Commuting from Car to Bicycle', *Science of The Total Environment*, 584/585, pp. 55-63

Kleiner, M. M. and Krueger, A. B. (2013), 'Analyzing the Extent and Influence of Occupational Licensing on the Labor Market', *Journal of Labor Economics*, 31(1), pp. 173-202

Lazear, E.P. (2000), 'Economic Imperialism', *Quarterly Journal of Economics*, 115 (1), pp. 99-146

Lazear, E. P. (2009), 'Firm-Specific Human Capital: A Skill-Weights Approach', *Journal of Political Economy*, 117 (5), pp. 914-940

Lazear, E. P. (2016), 'Overconfidence and Occupational Choice', NBER Working Papers, 21921

Lin, J. J.; Wang, N. L. and Feng, C. M. (2017), 'Public Bike System Pricing and Usage in Taipei', *International Journal of Sustainable Transportation*, 11, pp. 633-641

Loury, G.C. (1979): 'Market Structure and Innovation', *Quarterly Journal of Economics*, 93 (3), pp. 395-410

Maddison, A. (2001), *The World Economy: A Millenial Perspective*, OECD

Marx, K. (1844), A Contribution to the Critique of Hegel's Philosophy of Right, *Deutsch-Französische Jahrbücher*, February

Mateo-Babiano, I.; Bean, R.; Corcoran, J. and Pojani, D. (2016), 'How Does our Natural and Built Environment Affect the Use of Bicycle-Sharing?', *Transportation Research Part A*, 94, pp. 295-307

Matzarakis, A., Rutz, F., and Mayer, H. (2007), 'Modelling Radiation Fluxes in Simple and Complex Environments - Application of the RayMan Model', *International Journal of Biometeorology*, 51(4), pp. 323-334

Matzarakis, A., Rutz, F., and Mayer, H. (2010), 'Modelling Radiation Fluxes in Simple and Complex Environments: Basics of the RayMan Model', *International Journal of Biometeorology*, 54(2), pp. 131-139

McCloskey, D. N. (2006), *The Bourgeois Virtues: Ethics for an Age of Commerce*, The University of Chicago Press

McLanahan, Sara and Sandefur, G.D. (1994), *Growing up with a Single Parent: What Hurts, What Helps*, Harvard University Press

Mullainathan, S. and Shleifer, A. (2005), 'The Market for News', *American Economic Review*, 95(4), pp. 280-316

Mulligan, C.B. (1997), *Parental Priorities and Economic Inequality*, The University of Chicago Press

Nair, R.; Miller-Hooks, E.; Hampshire, R. C. and Busi, A. (2013), 'Large-Scale Vehicle Sharing Systems: Analysis of Vélib', *International Journal of Sustainable Transportation*, 7, pp. 85-106

Neal, D. (1995), Industry-Specific Human Capital: Evidence from Displaced Workers, *Journal of Labor Economics*, 13 (4), pp. 653-677

Nisbett, R.E. and Cohen, D. (1996), *Culture Of Honor: The Psychology Of Violence In The South*, Routledge

Noland, R. B.; Smart, M. J. and Guo, Z. (2016), 'Bikeshare Trip Generation in New York City', *Transportation Research Part A*, 94, pp. 164-181

O'Keeffe, A. G. and Baio, G. (2016), 'Approaches to the Estimation of the Local Average Treatment Effect in a Regression Discontinuity Design', *Scandinavian Journal of Statistics*, 43, pp. 978-995

Olafsson, A. S.; Nielsen, T. S. and Carstensen, T. A. (2016), 'Cycling in Multimodal Transport Behaviours: Exploring Modality Styles in the Danish Population', *Journal of Transport Geography*, 52, pp. 123-130

Park, C. and Sohn, S. Y. (2017), 'An Optimization Approach for the Placement of Bicycle-Sharing Stations to Reduce Short Car Trips: An Application to the City of Seoul', *Transportation Research Part A*, 105, pp. 154-166

Posner, R. A. (1973), *Economic Analysis of Law*, Boston: Little Brown, 1st edition.

Raux, C., Zoubir, A. and Geyik, M. (2017), 'Who are Bike-Sharing Schemes Members and do They Travel Differently? The Case of Lyon's "Velo'v" Scheme', *Transportation Research Part A*, 106, pp. 350-363

Ricci, M. (2015), 'Bike-sharing: A Review of Evidence on Impacts and Processes of Implementation and Operation', *Research in Transportation Business & Management*, 15, pp. 28-38

Rosen, S. (1981), The Economics of Superstars, *American Economic Review*, 71 (5), pp. 845-858

Saneinejad, S.; Roorda, M. J. and Kennedy, C. (2012), 'Modelling the Impact of Weather Conditions on Active Transportation Travel Behaviour', *Transportation Research Part D*, 17, pp. 129-137

Smith, A. (1904), *An Inquiry into the Nature and Causes of the Wealth of Nations*, ed.Edwin Cannan, London:Methuen, 1904

Stigler, G. J. (1966), *The Theory of Price*, New York: Macmillan, 3rd edition.

Stigler, G. J. (1971), 'The Theory of Economic Regulation', *The Bell Journal of Economics and Management Science*, 2(1), pp. 3-21

Sweeny, K.; Melnyk, D.; Miller, W. and Shepperd, J. A. (2010),'Information Avoidance: Who, What, When, and Why', *Review of General Psychology*, 14(4), pp. 340-353

Taylor, K. (2017), '40 of the Biggest Scandals in Uber's History', *Business Insider*, Nov. 24, 2017

Wang, X.; Lindsey, G.; Schoner, J. E. and Harrison, A. (2016), 'Modeling Bike Share Station Activity: Effects of Nearby Businesses and Jobs on Trips to and from Stations', *Journal of Urban Planning and and Development*, 142, pp. 1-9

Weber, Max (1905), *The Protestant Ethic and the Spirit of Capitalism*, New York : Scribner