Essays in Applied Microeconomics-Collection of Results

Submitted to the Doctoral School of Economics, Business and Informatics, Corvinus University of Budapest

Péter Isztin
supervised by Barna Bakó

September 15, 2021

1 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 Barna Bakó, 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.

2 Heads in the Sand:Information Aversion in a Market Context

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.

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. $e_i \in [0, 1]$. Specifically, we assume that the utility can be written in the following form

$$U_i(A, e_i) = v(A) - u(e_i) \tag{1}$$

where $u(\cdot)$ is continuous and increasing function, while $U'(e_i) < 0$ and $U''(e_i) \geq 0$.

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

$$\phi v(a) \geq e_i \tag{2}$$

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.

2.1 Laissez-faire

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

$$W^{LF} = \phi v(a) + \psi(Z)\Delta_1 - (1-\psi(Z)) \int_0^{\phi v(a)} e \text{ de} + \psi(Z)\left(\phi v(a) + \Delta_2\right) + (1-\psi(Z))(1-\phi v(a))\phi v(a) - IZ \tag{3}$$

and

$$CW^{LF} = \psi(Z)(\Delta_1 + \Delta_2) - (1-\psi(Z)) \int_0^{\phi v(a)} e \text{ de} \tag{4}$$

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 $t = 1, 2$.

2.2 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 favour 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

$$(1-\eta)\phi v(a) > s \tag{5}$$

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

$$\phi v(a) + (1-\eta)\phi v(a) - s \geq e_i \tag{6}$$

in case when (5) holds. On the other hand, if (5) is not satisfied, then consumers behave as in the case of laissez-faire and listen to new information as long as (2) holds. In what follows we assume that $(1-\eta)\phi v(a) > s$ holds.[1]

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

From (6) 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

$$E\pi = \phi v(a) - (1-\psi(Z))(1-\eta)e^*\phi v(a) + \psi(Z)\max\{e^*v(a), \phi v(a)\} + (1-\psi(Z))(1-e^*)\phi v(a) - IZ \quad (7)$$

Simplifying (7) yields

$$E\pi = \begin{cases} \phi v(a)[1-(1-\psi(Z))((2-\eta)e^*-1)] + \psi(Z)v(a)e^* - IZ & s < \phi[(2-\eta)v(a)-1] \\ \phi v(a)[1-(1-\psi(Z))((2-\eta)e^*-1)] + \psi(Z)v(a)\phi - IZ & s \geq \phi[(2-\eta)v(a)-1] \end{cases} \quad (8)$$

Maximizing (8) with respect to Z, we get the following first-order conditions

$$\psi'(Z)v(a)[(\phi((2-\eta)e^*-1)+e^*] = I \quad (9)$$

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

$$\psi'(Z)\phi v(a)(2-\eta)e^* = I \quad (10)$$

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:

$$W^{TL} = \phi v(a) + (1-\psi(Z))\left[(1-e^*)\phi v(a) - \int_0^{e^*} e \text{ de} - e^*s\right] - IZ + \psi(Z)\Delta_1 + \begin{cases} \psi(Z)e^*v(a) & s < \phi[(2-\eta)v(a)-1] \\ \psi(Z)(\phi v(a)+\Delta_2) & s \geq \phi[(2-\eta)v(a)-1] \end{cases} \quad (11)$$

while

$$CW^{TL} = \psi(Z)\Delta_1 + (1-\psi(Z))(1-\eta)e^*\phi v(a) - (1-\psi(Z))\left[\int_0^{e^*} e \text{ de} + e^*s\right] + \begin{cases} 0 & s < \phi[(2-\eta)v(a)-1] \\ \psi(Z)(\Delta_2) & s \geq \phi[(2-\eta)v(a)-1] \end{cases} \quad (12)$$

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

2.3 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.[2] 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.

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

$$E\pi = \psi(Z)(2\hat{\phi}v(a)) + (1 - \psi(Z))\lambda[\hat{\phi}v(a) + (1 - \hat{\phi}v(a))\hat{\phi}v(a)] - IZ \tag{13}$$

Taking the first derivative of equation (13) with respect to Z yields the following first-order condition

$$\psi'(Z)\hat{\phi}v(a)[2 - \lambda(2 - \hat{\phi}v(a))] = I \tag{14}$$

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 ϕ 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

$$W^R = \psi(Z)[2\hat{\phi}v(a) + \hat{\Delta}_1 + \hat{\Delta}_2] + (1 - \psi(Z))\lambda\left[\hat{\phi}v(a) + (1 - \hat{\phi}v(a))\hat{\phi}v(a) - \int_0^{\hat{\phi}v(a)} e \, \mathrm{d}e\right] - IZ \tag{15}$$

and

$$CW^R = \psi(Z)[\hat{\Delta}_1 + \hat{\Delta}_2] - (1 - \psi(Z))\lambda\left[\int_0^{\hat{\phi}v(a)} e \, \mathrm{d}e\right] \tag{16}$$

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

[2] 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.

2.4 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.

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

3.1 Introduction

This chapter assesses the impact of the exit of Uber from Budapest on BSS ridership.

3.2 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 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 (**?**) to mitigate the unobservable changes that might occur in a larger time window.

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 1: Number of trips for the summers of 2015 and 2016

Table 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.

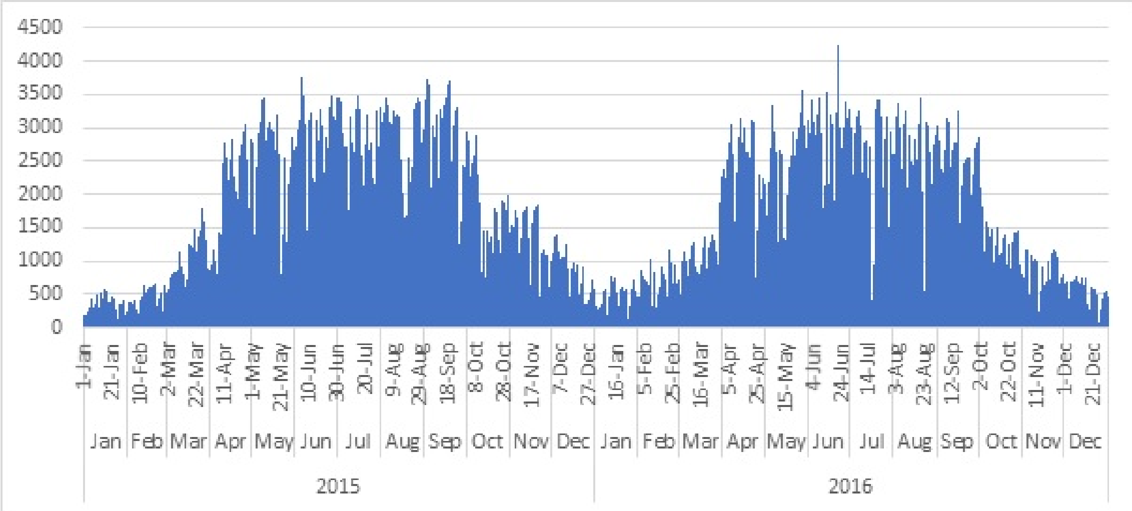

Figure 1: Daily usage frequencies of the Budapest BSS (total number of trips per day)

In this study, we exploit the fact that Uber 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.

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 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 1).

Our model can be written in the following general form:

$$\Delta y_{it} = \beta \Delta Uber_t + \Gamma \Delta x_{it} + c_i + u_{it} \tag{17}$$

3.3 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

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: Summary statistics

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 (17). Table 3 summarizes the 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 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 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), 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 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.

3.4 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

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 3: Estimation results

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).

Results are summarized in Table 4. Control variables were eliminated from the table to reduce its size. More detailed results are presented in the Appendix. 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

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 4: Effect of Uber on BSS usage by time periods

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.

3.5 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.

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.

4 A Theory of Early and Late Specialization

4.1 Introduction

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.

4.2 The Model

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.

The parent maximizes

$$\max_{\underline{H}, H_s} u(C_p) + a\delta \max\left\{(W_0 + \theta_1\mu_1)(H_1, H_t, \underline{H}); ...; (W_0 + \theta_s\mu_s)(H_S, \underline{H})\right\} \quad (18)$$

subject to the time constraint

$$T = C_p + \underline{h} + h_1 + ... + h_S + f \quad (19)$$

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

$$a\Bigg(\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s}+\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s}+ \frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\Bigg)=\lambda \tag{20}$$

as well as

$$a\Bigg(\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}}+\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}\Bigg)=\lambda. \tag{21}$$

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:

$$\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s}+\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s}+ \frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}= \frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}}+\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}, \tag{22}$$

$$\frac{\partial v(C_p)}{\partial C_p}=a\Bigg(\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_s}\frac{dH_s}{dh_s}+ \frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s}+\frac{\partial(W_0+\max\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\Bigg) \tag{23}$$

and

$$\frac{\partial v(C_p)}{\partial C_p}=a\Bigg(\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}}+\frac{\partial(W_0+\sum_S\theta_s\mu_s)(H_s,H_t,\underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}\Bigg). \tag{24}$$

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.

4.3 Castes and hereditary positions

4.3.1 "Separate but equal"

First we consider a case where there are no status doifferences across the castes. Assume there are n individuals, half of which belong to one demographic group and the other half to an other demographic group. In each group, half of the individuals have a comparative advantage in task A, while the other half has a comparative advantage in task B. Assume the market is in equilibrium when the number of workers in each market is $\frac{N}{2}$. Now let us write down the parnet's problem.

Her objective function in a free market system is

$$u(C_p) + a\Big(\frac{1}{2}(W_0 + \mu_A)(H_A, H_t, \underline{H}) + \frac{1}{2}(W_0 + \mu_B)(H_B, H_t, \underline{H}, H_t)\Big). \qquad (25)$$

The first order conditions are

$$a\frac{1}{2}\bigg(\frac{\partial(W_0 + \mu_s)(H_s, H_t, \underline{H})}{\partial H_s}\frac{dH_s}{dh_s} + \frac{\partial(W_0 + \mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s} + \frac{\partial(W_0 + \mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\bigg) = \lambda \qquad (26)$$

for specific skills, as well as

$$a\bigg(\frac{\partial(W_0 + \mu_s)(H_s, H_t, \underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}} + \frac{\partial(W_0 + \mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}\bigg) = \lambda \qquad (27)$$

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

$$a\bigg(\frac{\partial(W_0 + \frac{1}{2}\mu_s)(H_s, H_t, \underline{H})}{\partial H_s}\frac{dH_s}{dh_s} + \frac{\partial(W_0 + \frac{1}{2}\mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial h_s} + \frac{\partial(W_0 + \frac{1}{2}\mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial H_s}\frac{dH_s}{dh_s}\bigg) = \lambda \qquad (28)$$

for specific human capital, as well as

$$a\bigg(\frac{\partial(W_0 + \frac{1}{2}\mu_s)(H_s, H_t, \underline{H})}{\partial \underline{H}}\frac{d\underline{H}}{d\underline{h}} + \frac{\partial(W_0 + \frac{1}{2}\mu_s)(H_s, H_t, \underline{H})}{\partial H_t}\frac{\partial H_t}{\partial \underline{h}}\bigg) = \lambda \qquad (29)$$

for general human capital.

In a free economy, both markets are served by those with a comparative advantage in that sector. In parents face uncertainty over comparative advantage, they invest relatively less in specific, but relatively more in general human capital, compared to a caste system, where they are guaranteed a place in one sector, however, they may be among those with a comparative disadvantage (and hence those without a rent earning capacity) in the sector.

4.3.2 Hierarchical castes

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. The goods produced by the sectors may be net substitutes, complements, or there may be a Cobb-Douglas preference relation between them. 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. First we consider a very general case. 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

$$W_I = (W_0 + \mu_I)(H_t, H_I, \underline{H}), \tag{30}$$

while in sector II it can be written as

$$W_{II} = (W_0 - \mu_{II})(H_t, H_{II}, \underline{H}). \tag{31}$$

The first order conditions are

$$a(W_0 + \mu_I)\frac{\partial H_I}{\partial h_I}\frac{\partial H_I}{\partial a_I}\left(\frac{\partial W_I}{\partial H_I} + \frac{\partial W_I}{\partial H_t} + \frac{\partial W_I}{\partial H_t}\frac{\partial H_t}{\partial H_I}\right) = \lambda, \tag{32}$$

$$a(W_0 + \mu_I)\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_I}{\partial \underline{H}} + \frac{\partial W_I}{\partial H_t}\right) = \lambda, \tag{33}$$

$$a(W_0 - \mu_{II})\frac{\partial H_{II}}{\partial h_{II}}\frac{\partial H_{II}}{\partial a_{II}}\left(\frac{\partial W_{II}}{\partial H_{II}} + \frac{\partial W_{II}}{\partial H_t}\frac{\partial W_{II}}{\partial H_t}\frac{\partial H_t}{\partial H_{II}}\right) = \lambda, \tag{34}$$

and

$$a(W_0 - \mu_{II})\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_{II}}{\partial \underline{H}} + \frac{\partial W_{II}}{\partial H_t}\right) = \lambda. \tag{35}$$

Combining equations (16) and (17) we obtain

$$\frac{\partial H_I}{\partial h_I}\frac{\partial H_I}{\partial a_I}\left(\frac{\partial W_I}{\partial H_I} + \frac{\partial W_I}{\partial H_t} + \frac{\partial W_I}{\partial H_t}\frac{\partial H_t}{\partial H_I}\right) = \frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_I}{\partial \underline{H}} + \frac{\partial W_I}{\partial H_t}\right) \tag{36}$$

as well as

$$\frac{\partial H_{II}}{\partial h_{II}}\frac{\partial H_{II}}{\partial a_{II}}\left(\frac{\partial W_{II}}{\partial H_{II}} + \frac{\partial W_{II}}{\partial H_t}\frac{\partial W_{II}}{\partial H_t}\frac{\partial H_t}{\partial H_{II}}\right) = \frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_{II}}{\partial \underline{H}} + \frac{\partial W_{II}}{\partial H_t}\right). \tag{37}$$

Who belongs to a privileged "class" is not always clear. Consider a less extreme version of our model: as above, there are two sectors, with barrier to entry in sector I. 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.

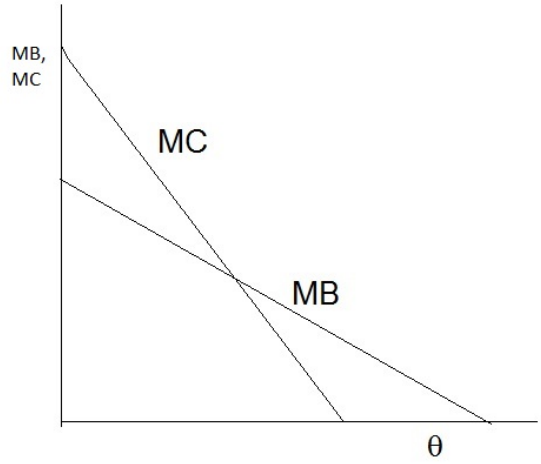

Figure 2: Marginal benefit and cost of "tightening the caste system"

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 3.1.

[3]More formally, the marginal benefit of an increase in θ can be expressed as $(W_0 + \mu_s)(H_t, H_s, \underline{H}) + \frac{dH_s}{d\theta} + \frac{d\underline{H}}{d\theta}$ 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 $\gamma'(\theta) < 0$.

4.4 Gender roles and the sexual division of labor

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 adult child's share in the marital surplus α and the probability that this market premium is realized, θ_h.

The expected "wage" in the household sector is

$$W_h = A_h \theta_h \alpha \mu(H_t, H_h, \underline{H}), \tag{38}$$

while in the market sector it can be written as

$$W_m = (1 - \theta_h) A_m \alpha W(H_t, H_m, \underline{H}). \tag{39}$$

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

$$A_h \alpha \theta_h \mu \frac{dH_h}{dh_h} \left(\frac{\partial W_h}{\partial H_h} + \frac{\partial W_h}{\partial H_t} + \frac{\partial W_h}{\partial H_t} \frac{\partial H_t}{\partial H_h} \right) = \lambda. \tag{40}$$

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

$$(1 - \theta_h) A_m \alpha W \frac{dH_m}{dh_m} \left(\frac{\partial W_m}{\partial H_m} + \frac{\partial W_m}{\partial H_t} \frac{\partial W_m}{\partial H_t} \frac{\partial H_t}{\partial H_m} \right) = \lambda. \tag{41}$$

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

$$\theta_h A_h \alpha \mu \frac{\partial \underline{H}}{\partial \underline{h}} \left(\frac{\partial W_h}{\partial \underline{H}} + \frac{\partial W_h}{\partial H_t} \right) + (1 - \theta_h) A_m W \frac{\partial \underline{H}}{\partial \underline{h}} \left(\frac{\partial W_m}{\partial \underline{H}} + \frac{\partial W_m}{\partial H_t} \right) = \lambda. \tag{42}$$

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

$$\begin{aligned} A_h \theta_h \mu \frac{dH_h}{dh_h} \left(\frac{\partial W_h}{\partial H_h} + \frac{\partial W_h}{\partial H_t} + \frac{\partial W_h}{\partial H_t} \frac{\partial H_t}{\partial H_h} \right) &= \\ \theta_h A_h \mu \frac{d\underline{H}}{d\underline{h}} \left(\frac{\partial W_h}{\partial \underline{H}} + \frac{\partial W_h}{\partial H_t} \right) + (1 - \theta_h) A_m W \frac{d\underline{H}}{d\underline{h}} \left(\frac{\partial W_m}{\partial \underline{H}} + \frac{\partial W_m}{\partial H_t} \right) & \end{aligned} \tag{43}$$

for household investments, and

$$
\begin{aligned}
(1-\theta_h)A_m W\frac{dH_m}{dh_m}\left(\frac{\partial W_m}{\partial H_m}+\frac{\partial W_m}{\partial H_t}\frac{\partial W_m}{\partial H_t}\frac{\partial H_t}{\partial H_m}\right) = \\
\theta_h A_h \mu\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_h}{\partial \underline{H}}+\frac{\partial W_h}{\partial H_t}\right)+(1-\theta_h)A_m W\frac{d\underline{H}}{d\underline{h}}\left(\frac{\partial W_m}{\partial \underline{H}}+\frac{\partial W_m}{\partial H_t}\right)
\end{aligned} \tag{44}
$$

for market investments.

4.5 Conclusions

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.

5 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.

5.1 Parent-child interaction

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. . 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. 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.

5.2 Santa Claus as a solution

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 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.

5.3 Religious persuasion in politics

5.4 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 with probability η^{n-1}, where n is the number of citizens. 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 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

$$V_c = w((1-\tau)x_i - \eta^{n-1}y_{-i} + \eta y_i) - ((1-\eta)f + \sqrt{g})y_i - {x_i}^2 - {y_i}^2 \tag{45}$$

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.

The optimal values for each are

$$x_i = \frac{(1-\tau)w}{2} \tag{46}$$

and

$$y_i = \frac{\eta(w+f) - f - \sqrt{g}}{2}. \tag{47}$$

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

$$R = \tau n w \left(\frac{(1-\tau)w}{2} - \eta \left(\frac{\eta(w+f) - f - \sqrt{g}}{2} \right) \right). \tag{48}$$

Maximizing with respect to τ we obtain the optimal tax rate

$$\tau = \frac{\eta(f + \sqrt{g}) + w - \eta^2(f+w)}{2w}. \tag{49}$$

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

$$\frac{\alpha n w(\eta(f + \sqrt{g}) + w - \eta^2(f+w))}{4} - c_g g - c_s s. \tag{50}$$

The church maximizes the objective function with respect to g and s.

The optimal level of g is given as

$$g = \frac{(\alpha \eta n w)^2}{64 {c_g}^2}. \tag{51}$$

We can also solve for the optimal tax rate, as well as the optimal x_i and y_i. We obtain

$$\tau = \frac{8w + \eta \left(8f + \sqrt{\frac{(\alpha\eta n w)^2}{{c_g}^2}} - 8\eta(f+w) \right)}{16w}, \tag{52}$$

$$x_1 = \frac{1}{32} \left(8w + 8\eta^2(f+w) - \eta \left(8f + \sqrt{\frac{(\alpha\eta n w)^2}{{c_g}^2}} \right) \right), \tag{53}$$

and

$$y_1 = \frac{1}{2} \left(f(\eta - 1) + \eta w - \frac{1}{8} \sqrt{\frac{(\alpha\eta n w)^2}{{c_g}^2}} \right). \tag{54}$$

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

$$\frac{8w + \eta\left(8f + \sqrt{\frac{(\alpha\eta n w)^2}{{c_g}^2}} - 8\eta(f+w) \right)}{16w} n w \left(\frac{1}{32}(8w + 8\eta^2(f+w) - \eta\left(8f + \sqrt{\frac{(\alpha\eta n w)^2}{{c_g}^2}} \right)) - \eta \frac{1}{2}\left(f(\eta-1) + \eta w - \frac{1}{8}\sqrt{\frac{(\alpha\eta n w)^2}{{c_g}^2}} \right) \right). \tag{55}$$

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

$$V_c = w((1-\tau)x_i - \eta^{n-1}y_{-i} + \eta y_i) - (1-\eta)fy_i - {x_i}^2 - {y_i}^2, \tag{56}$$

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

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

$$x_i = \frac{(1-\tau)w}{2}. \tag{57}$$

while the optimal y_i is

$$y_i = \frac{\eta(w+f)-f}{2}. \tag{58}$$

The leader now maximizes

$$R = \tau nw\left(\frac{(1-\tau)w}{2} - \eta\left(\frac{\eta(w+f)-f}{2}\right)\right). \tag{59}$$

Maximizing with respect to τ we obtain the optimal tax rate

$$\tau = \frac{\eta(f-\eta(f+w))+w}{2w}. \tag{60}$$

Plugging the equilibrium values into the tax revenue we obtain

$$\frac{\eta(f-\eta(f+w))+w}{2w}nw\left(\frac{(1-\tau)w}{2} - \eta\left(\frac{\eta(w+f)-f}{2}\right)\right). \tag{61}$$

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

$$n \geq \frac{8\left(\sqrt{\frac{1}{1-\alpha}}-1\right)c_g(1-\eta)(w+\eta(f+w))}{\alpha\eta^2 w} \tag{62}$$

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."

5.5 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.

Toward an Economics of Moral Character

6 Introduction

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.

6.1 Character in consumption

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 and third period 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). Crucially, the wage is also a function of y and V.

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:

$$\frac{\partial U}{\partial x} = \lambda p_x, \tag{63}$$

and

$$\frac{\partial U}{\partial y} + \frac{\partial U}{\partial S}\frac{dS}{dy} = \lambda p_y + \int_0^t e^{-\rho t} p_y. \tag{64}$$

These first-order conditions establish an optimal x and y, x^* and y^*. The consumption of the numeraire is a function of the lifetime wealth, 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

$$V_p = U(C_p) + aV_c(x^*(W_c(S)), y^*(W_c(S), p_y, S, V, \beta(V))). \tag{65}$$

The first order conditions in the steady state are

$$\frac{dU(C_p)}{dC_p} = \lambda \tag{66}$$

and

$$a\frac{dV}{dv}\left(\frac{\partial x^*}{\partial W_c}\frac{\partial W_c}{\partial S}\frac{\partial S}{\partial y}\frac{\partial y}{\partial V} + \frac{\partial V_c}{\partial y^*}\frac{\partial y^*}{\partial V} + \frac{\partial V_c}{\partial S}\frac{dS}{dy}\left(\frac{\partial y}{\partial V} + \frac{\partial y}{\partial \beta}\frac{\partial \beta}{\partial V}\right)\right) = \lambda. \tag{67}$$

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.

6.2 Fully altruistic model with transfers and human capital investments

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.

6.2.1 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

$$V_p{}^y(C_p{}^y) + V_p{}^m(C_p{}^m) + V_p{}^o(C_p{}^o) + a(V_c{}^y + V_c{}^m), \tag{68}$$

subject to the intertemporal budget constraint

$$c_p{}^y + \frac{c_p{}^m}{1+r} + \frac{c_p{}^o}{(1+r)^2} + \frac{t}{(1+r)^2} + h + v = m_p{}^y + \frac{m_p{}^m}{1+r} + \frac{m_p{}^o}{(1+r)^2}. \tag{69}$$

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

$$a\frac{\partial V_c{}^m}{\partial t} = \lambda \tag{70}$$

and

$$\frac{dV_p{}^y}{dC_p{}^y} + \frac{dV_p{}^m}{dC_p{}^m} + \frac{dV_p{}^o}{dC_p{}^o} = \lambda, \tag{71}$$

which implies

$$a\frac{\partial V_c{}^m}{\partial t} = \frac{dV_p{}^y}{dC_p{}^y} + \frac{dV_p{}^m}{dC_p{}^m} + \frac{dV_p{}^o}{dC_p{}^o}. \tag{72}$$

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.

6.2.2 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. Resouces 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

$$a\frac{\partial V_c}{\partial H_c}\frac{dH_c}{dh} = \lambda, \tag{73}$$

while the FOC with respect to own consumption is

$$\frac{d{V_p}^y}{d{C_p}^y} + \frac{d{V_p}^m}{d{C_p}^m} + \frac{d{V_p}^o}{d{C_p}^o} = \lambda. \tag{74}$$

This implies

$$a\frac{\partial V_c}{\partial H_c}\frac{dH_c}{dh} = \frac{d{V_p}^y}{d{C_p}^y} + \frac{d{V_p}^m}{d{C_p}^m} + \frac{d{V_p}^o}{d{C_p}^o}. \tag{75}$$

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

$$\begin{aligned} &a\frac{dV}{dv}\left(\beta\frac{\partial x^*}{\partial W_c}\frac{\partial W_c}{\partial S}\frac{dS}{dy}\frac{\partial y}{\partial V} + \frac{\partial V_c}{\partial y^*}\frac{\partial y^*}{\partial V} + \beta\frac{\partial V_c}{\partial S}\frac{dS}{dy}\left(\frac{\partial y}{\partial V} + \beta\frac{\partial y}{\partial \beta}\frac{\partial \beta}{\partial V}\right)\right) + \\ &\quad \frac{d{V_p}^y}{d{C_p}^y} + \frac{d{V_p}^m}{d{C_p}^m} + \frac{\partial {C_c}^m}{\partial t}\frac{\partial t}{\partial V}\left(\frac{d{{V_p}^o}^o}{d{C_p}^o}\frac{\partial {C_p}^o}{\partial V_c}\frac{\partial V_c}{\partial t}\frac{\partial t}{\partial V}\right) - \\ &\quad a\frac{dV}{dv}\left(\frac{\partial V_c}{\partial t}\frac{\partial t}{\partial V} + \frac{\partial V_c}{\partial H_c}\frac{dH_c}{dh}\frac{\partial h}{\partial^2 V}\right) = \lambda. \end{aligned} \tag{76}$$

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 o the parent's part. This in turn decreases the amount of money transferred by the parent to the adult hild. This directly increases the consumption of the parent but decreases the utility of the child.

6.3 Crime and punishment

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

$$v(y) + b(x) - \beta p f(x) - m(x), \tag{77}$$

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

$$b'(x) = pf'(x) + m'(x), \tag{78}$$

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

$$v(y) + b(x^*(\beta, p, f, m)) - \beta p f(x^*(p, f, m)) - m(x^*(p, f, m)). \tag{79}$$

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

$$v(C_p) + a(v(y) + b(x^*(\beta, p, f, m)) - pf(x^*(\beta, p, f, m)) - m(x^*(\beta, p, f, m), V)). \tag{80}$$

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

$$\frac{dV}{dv}\left(\frac{\partial b}{\partial x^*}\frac{\partial x^*}{\partial m}\frac{\partial m}{\partial V} - p\frac{\partial f}{\partial x^*}\frac{\partial x^*}{\partial m}\frac{\partial m}{\partial V} - \frac{\partial m}{\partial x^*}\frac{\partial x^*}{\partial m}\frac{\partial m}{\partial V} + p\frac{\partial f}{\partial x^*}\frac{\partial x^*}{\partial \beta}\frac{\partial \beta}{\partial v}\right) = \frac{\partial g}{\partial v}. \tag{81}$$

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.

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

$$W_i(t_i) + W_l(h, t_l) - \beta pf(t_i) - \beta m(t_i). \quad (82)$$

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

$$\begin{aligned} V_c = W_i(t_i(\beta, p, f, m)) + W_l(h, t_l(\beta, p, f, m, h)) - \\ \beta(pf(t_i(\beta, p, f, m)) + m(V, t_i(\beta, p, f, m))). \end{aligned} \quad (83)$$

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

$$\begin{aligned} V_p = u(C_p) + a[W_i(t_i(W_l(h), \beta, p, f, m(V))) + W_l(h, t_l(W_l(h), \beta, p, f, m(V), h)) - \\ \beta(pf(t_i(, W_l(h), \beta, p, f, m(V))) + m(v, t_i(W_l(h), \beta, p, f, m(V))))]. \end{aligned} \quad (84)$$

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

$$\frac{\partial W_l}{\partial h} t_l = \lambda. \quad (85)$$

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.

6.4 Moral character 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.

Now 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

$$pW - e \geq qW - m. \tag{86}$$

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

$$W^* = \frac{e - m}{p - q}. \tag{87}$$

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.

6.5 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.

7 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.

References

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

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