Market Behavior and Market Interaction under Incomplete Information: Theory and Empirical Implications

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Wolfram Merzyn
aus Göttingen.

Tag der Promotion: 15.06.2007

Diese Disseration ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Dekan:	Prof. Dr. Gerhard Wagner
Erstreferent:	Prof. Dr. Georg Nöldeke
Zweitreferent:	Prof. Dr. Benny Moldovanu

Tag der Promotion: 15.06.2007

Contents

Introduction

All three chapters of this dissertation deal with the question of how individuals behave and interact in markets. Nevertheless, they are only loosely related, everyone of them exploring a very different aspect of the same broad theme. For this reason, they have been designed to form independent and self-contained units. The main purpose of this introduction is to give potential readers an idea of what they can expect from reading the subsequent chapters at full length.

Chapter 1 takes the following observation as its starting point: In many cases, an individual's willingness to pay for a certain unit of a good does not only depend on her valuation of the good, but also on what she thinks its going price is. In order to understand why this statement should be true, let us consider the following example: Imagine an employee who wants to buy a bike for her daily way to work. To make matters stark, let us assume that the employee will lose her job if she does not get a bike. May this dependency lead her to accept straightaway if she is offered a bike at the price of a car? The answer to this question is "no", because the employee will feel very confident that a much better deal is available at the next bike store. Hence, there is no reason for her to behave as if the only alternative to buying the overpriced bike was not to buy any bike at all. This simple insight applies not only to bikes, but to all goods for which a sufficiently liquid market exists.

Assume next that the employee from our example only has a vague idea of how expensive bikes usually are. It is probably clear to her that bikes are much cheaper than cars, but she may still be uncertain about whether a given price is relatively high or low. One way to overcome this lack of information - the one we focus on - is to pay attention to the behavior of her "competitors" (i.e. other potential buyers). If, for example, a given bike sells within minutes after it has first been offered, she can conclude that the bike most likely was relatively cheap. This piece of information allows her to form a more accurate belief about what a low price is and adjust her willingness to pay accordingly. At the same time, her own behavior serves as a signal to other potential buyers.

Our goal in the first chapter is to conduct a formal analysis of how potential buyers behave and interact in the presence of uncertainty about the distribution of prices. To this end, we consider a model of a big, decentralized auction market (akin to eBay). Two types of agents are active in this market: buyers and sellers. Buyers want to buy one unit of the good under consideration, while sellers have one to offer. Buyers can successively participate in more than one auction. After every auction, however, they face some probability of "dying", so there is pressure on them to win an auction as soon as possible. The idiosyncratic feature of our model is that no buyer knows how many other buyers there are. Hence, there is some uncertainty about how intensely the buyers compete against each other. The auctions in our model are assumed to be of the second-price, sealed-bid type, and the buyers cannot directly observe each other's behavior, but if they lose an auction, they can conclude that the highest competing must have been higher than their own one. This insight allows them to update their belief about the "state of the world".

We characterize the equilibrium behavior of the buyers in our model. Our main result is that the buyers in our dynamic model behave the same as the buyers in a standard (static) common value auction (i.e. an auction for a good with unknown characteristics). Thus, much of what is known about common value auctions also applies to environments in which goods with *known* characteristics are for sale, but buyers do not know the distribution of winning bids.

In **Chapter 2** we analyze the effect of redistribution on the demand for "status goods". Doing so requires us to first make precise what status goods are and why there is demand for them. We assume that individuals want to be perceived as attractive. No individual, however, has an opportunity to directly show off her qualities. Yet, there is a positive correlation between a person's attractiveness and her income. The income itself is unobservable as well, but an individual can publicly "burn" part of it so as to let everyone know how rich - and, hence, attractive - she is. Consuming status goods (like, for example, Rolex watches) is one way of publicly burning money. (In the model we consider it is the only one.) How does the positive relationship between affluence and attractiveness come about? In our model it is not true that a person's attractiveness directly depends on her income. Rather, the causal relationship runs the other way round: We assume that attractive (i.e. handsome, intelligent, eloquent, entertaining) individuals make more money than others. This assumption is important, as it implies that redistributing income does not affect the distribution of attractiveness in the population. Even if all differences in income were eliminated, some individuals would remain more attractive than others.

Our analysis reveals that redistribution has two contrarian effects on the demand for status goods: On the one hand, redistribution (from the rich to the poor) increases the demand for status goods, as it makes it more difficult for the rich to set themselves apart from the poor. This result is very intuitive: In a society with a very unequal income distribution, the rich need to spend only a small share of their income on status goods in order to make clear "who is who". The poor do not have any chance to behave as if they were rich.

There is, however, a second effect which works in precisely the opposite direction. This second effect is based on the assumption that an individual's propensity to consume the status good does not only depend on her income, but also on some other factors like, for example, personal preferences. Redistribution increases the importance of these "noise" factors, thereby reducing the reliability of consumption patterns as signals of attractiveness. Again, this effect is intuitive: If differences in income are largely taxed away, even individuals with modest gross income can afford a Rolex if they make this a high priority. Conversely, the rich are not as rich as they used to be, so they may well do without a Rolex if they do not care much about how they are perceived by others anyway. Therefore, we cannot be sure anymore that whoever owns a Rolex must be a member of the upper class. Everyone knows this, so there is no reason to buy a Rolex. While this mitigating effect matters little when income differences are large, we demonstrate its power in egalitarian societies. Hence, the consumption of status goods is highest at intermediate levels of income inequality.

In **Chapter 3** of this thesis we turn to the question of whether (and, if so, how) the information contained in individual schooling decisions can help us to estimate the so-called "returns to schooling", i.e. the causal effect of schooling on earnings. In order to understand why this is an interesting question, we need to reach back into the history of micro-econometric thought: The traditional and obvious way of estimating the returns to schooling is to compare the wages of individuals with different schooling achievements. There is, however, a major problem associated with this approach: as people, rather than being randomly assigned a certain level of schooling, decide themselves when to enter the labor market, individuals with high schooling attainments systematically differ from those with low ones. Hence, it is not clear what accounts for the observed positive relationship between schooling and earnings. In econometric terms the problem is said to lie in the "endogeneity"of the schooling variable.

One way to overcome this problem is to look for an observable "random component" in individual schooling decisions, i.e. a variable which has an influence on the timing of school exit, but is unrelated to any student's personal characteristics. If such a variable is available, one can estimate the returns to schooling by essentially comparing the earnings of students who have different schooling attainments, but do not systematically differ otherwise. Finding an observable random component in schooling decisions, however, is usually very difficult.

For this reason, we examine an alternative way of getting around the endogeneity problem, advocated by Belzil and Hansen (2002). The basic idea behind their approach is to exploit the information that is contained in individual schooling decisions. Intuitively, this can be done as follows: If an individual incurs certain observable costs in order to (voluntarily) go to school, we learn something about the returns to schooling: These must be at least as high as the costs, for otherwise the individual would not have chosen to go to school. Conversely, every decision to exit school indicates that the costs of remaining in school would have been higher than the benefits.

While the identification strategy we have just sketched sounds appealing, we show it to have several important weaknesses. In particular, the strategy only works if the costs of going to school are observable. Since going to school has a number of non-monetary costs and benefits, this is generally not the case. In fact, even the monetary costs of going to school are often difficult to observe. This insight leads us to investigate the way in which Belzil and Hansen identify the returns to schooling. We show that their identification strategy crucially relies on the assumption that individual schooling decisions are partly driven by (observable) random wage shocks. Thus, while the approach of Belzil and Hansen is an interesting one, it has much more in common with the traditional "random components" method than one might think at first glance.

Chapter 1

Sequential Common Value Auctions with Purely Private Valuations

1.1 Introduction

Over the past two decades the common value (or "mineral rights") model has become one of the most commonly used tools in auction theory. Its popularity mainly stems from two sources: First, the model is mathematically convenient and makes for a number of elegant theorems. Second, its applicability is widely believed to go beyond auctions for drilling rights or industrial procurement. More specifically, many economists hold that the common value (CV) model provides a useful description of auctions with resale opportunities[1] and, conversely, interpret observed common value elements as evidence of resale-motivated buyers.[2] This view, however, is largely based on simple reinterpretations of the standard CV model, i.e. on analyses which take the resale value of a good as exogenously given. Studying a model in which that value is *endogenously* determined, Haile (2003) does not only conclude that "standard models will often miss important aspects of bidding strategies" (in auctions with resale), but

[1]Examples include McAfee and McMillan (1987), Milgrom (1989), and Crampton (1995).

[2]See, for example, Bajari and Hortacsu (2003).

also remarks that "resale does not always create common or affiliated values".

In this chapter we consider a dynamic equilibrium model of an auction market to study a phenomenon closely related to, but still distinct from, auctions with resale. Rather than assuming that the winner of an auction has a chance to resell the good, we allow the losers to bid for an identical item in another auction, i.e. we assume that the item currently for sale has substitutes. As others have noted before us, the buyers' optimal strategy in such a setting is to *shade* their bids, i.e. to account for the possibility of future gains (in case of a loss) by bidding less than they would do in the absence of any alternative opportunities.[3] More precisely, the buyers should simply subtract the opportunity cost of winning from their valuation. In the case of a second-price sealed-bid auction, the optimal bidding function then takes the form $b = v - V$, where V denotes the value of bidding in future auctions.

Novel to our paper is the assumption that the buyers are *imperfectly informed* about how valuable the option of participating in future auctions is, because they do not know how scarce the good under consideration is. The main result of our paper is that this uncertainty transforms otherwise private value auctions into standard *common* value auctions. Thus, the failure of the CV model to accurately describe auctions with resale does not extend to the seemingly similar case of sequential auctions with aggregate uncertainty.[4]

Since both assumptions - substitutability and information imperfectness - seem to be rather generic, our results make a strong case for the CV paradigm. In particular, non-negligible common value elements can be expected to be the norm (rather than an exception) in most online-auctions where similar or even identical items are often sold at high frequency. This may help to explain the empirical findings of Bajari and Hortacsu (2003) and Roth and Ockenfels (2002) who discover evidence of common

[3] See, for example, Milgrom and Weber (1999).

[4] The term "aggregate uncertainty" (as opposed to "idiosyncratic uncertainty") refers to uncertainty about the *distribution* of buyer's characteristics (rather than uncertainty about any agent's *individual* characteristics).

values in data from auctions for goods with known characteristics like computers (i.e. "private value goods").

To be sure, the idea that aggregate uncertainty may induce a common value element to an otherwise "private value" trading environment is not entirely new to the literature[5]. So far, however, it has never been scrutinized, i.e. there does not yet exist a dynamic model of decentralized trading in which a common value component arises out of buyers' uncertainty about future prices (i.e. highest competing bids) and in which the distribution of prices is an *equilibrium object* of the model. The purpose of this paper is to fill this gap.

The structure of our model heavily borrows from the literature on dynamic matching and bargaining games:[6] There is an infinite sequence of time periods, and in every one of them continuums of buyers and sellers are born into the model. All buyers (the new ones and those inherited from the previous period) then get randomly and independently matched to one of the sellers. Every seller conducts a second-price, sealed bid auction. After the auction she gives notice to the winner and tells her how much to pay. Other than that, however, no buyer gets to know anything about the number or bids of her respective opponents. At the end of each round all successful agents (i.e. those buyers who managed to buy an item and the sellers who sold theirs) leave the model. All other agents can try their luck again in the next round unless they get "discouraged" which happens with exogenously given probability $1-\delta$. While the sellers do not exhibit any strategic behavior (they do not take any decisions), the buyers aim at maximizing expected utility which is their valuation minus the payment if they manage to win an auction and zero if they get discouraged.

The defining feature of our model is that the mass of inflowing buyers can (depending on the "state of the world") be either large or small; it does not change over time,

[5] Serrano (1995), for example, refers to aggregate uncertainty as "common value uncertainty".

[6] Important references include Gale (1987), Wolinsky (1990), McAfee (1993), and Satterthwaite and Shneyerov (forthcoming). The seminal paper on dynamic matching and bargaining under aggregate uncertainty is Wolinsky (1990).

though. Competition among buyers is more intense in the high state, so the "continuation value" of participating in future auctions is lower. Therefore, the buyers would like to bid more in the high state than in the low one. They are, however, assumed not to know the state in which they are. Upon birth, every buyer forms an initial belief about the state of the world. Moreover, after every auction the losing buyers get additional information regarding the state, because they are able to make some inference about the bidding behavior of their opponents. Thus, the buyers' beliefs are endogenous and evolve over time. A common value component arises in the auctions of our model, because the value a buyer puts on winning an auction depends on her opponents' private information about the state of the world.

In our analysis of the model, we focus on equilibrium *steady states*, i.e. on distributions of buyers' beliefs which do not change over time, because the buyers' learning process is just evened out by the turnover among them. What we demonstrate below is that an equilibrium pair of steady-state belief distributions (one for each state of the world) exists and that it is unique within a certain class of equilibria.

In contrast to many other models of dynamic auctions, we do not assume the number of buyers and sellers active in the market to be finite i.e. our analysis mainly applies to large auction markets like ebay. Consequently, the probability of any two agents meeting more than once is zero, so no buyer can possibly gain from deliberately manipulating other buyers' beliefs (using her bid as a signal). This simplification allows us to focus exclusively on the role of buyers' uncertainty about the state of the world.

We argue that our model can be used to analyze a wide range of applied questions. One of them is whether the revenue of a seller would change if she revealed the size of some bids or the number of bidders after her auction (and the buyers knew this beforehand). We show that, for any given distribution of buyers' types, announcing the winning bid has an adverse affect on revenue. The intuition behind this result is that any (ex-post) revelation of information only serves to increase the buyers' continuation payoff, thereby lowering all bids.[7]

The chapter is organized as follows: In Section 1.2 we formalize the ideas laid out in this introduction by setting up a model of sequential auctions with uncertainty about the ratio of buyers to sellers in the market ("aggregate demand uncertainty"). In Section 1.3 we show that this model has a unique steady state equilibrium and characterize some of its properties. Section 1.4 asks whether sellers could increase their revenue by employing a more transparent informational regime than the one assumed in Section 1.3. Finally, a discussion of the model and some conclusions are presented in Section 1.5.

1.2 The Model

1.2.1 Basics

The model we consider is set up in discrete time going from $t = -\infty$ to $t = \infty$, and the arena in which all of its action takes place is a big, decentralized market place (akin to a bazar) at which an indivisible, homogeneous good is traded. There are continuums of two types of agents active in the market: buyers and sellers. Every seller offers one unit of the good for sale, while every buyer seeks to purchase one. The timing within every period is as follows: At the beginning of every period exogenously given masses of buyers and sellers are born into the model. All buyers (new and old ones) are then independently and randomly matched to one of the sellers. The sellers allocate their goods by means of second-price sealed-bid auctions. At the end of the period all agents who have managed to trade leave the market. In addition, all other agents face some risk of being forced to leave. Technically, each of them is removed with exogenously given probability $1 - \delta$. The purpose of δ is to make the agents "impatient", i.e. its role is similar to that of a discount factor.[8]

[7] This result is reminiscent of the findings of Mezzetti, Pekec, and Tsetlin (2004) who make a similar point in the framework of a more stylized, two-period model.

[8] While we assume δ to be the same for both buyers and sellers, that assumption could easily be relaxed.

1.2.2 The Sellers and their Problem

The sellers do not face any decision problem at all, i.e. their role is that of an "auction running robot". In particular, we do not allow the sellers to set a reserve price. This assumption is made for simplicity only and does not drive any of the results in the paper. Nevertheless, it is likely to be perceived as restrictive. Therefore, we provide an extensive discussion in Section 1.5.

1.2.3 The Buyers and their Problem

The mass of sellers who enter the market every period is known and normalized to one. The mass of incoming buyers (denoted as d), on the other hand, can, depending on the state of the world, be either large ($d = d_h$) or small ($d = d_l$). The state does not change throughout. Both states are (ex ante) equally likely to occur, and the buyers do not know in which one of them they are. Thus, one important characteristic of every buyer is the probability she attaches to being in the high state, denoted as θ. Upon birth, all buyers receive two pieces of information about the state of the world: First, the mere fact of being alive allows them to conclude that - leaving other sources of information aside - the probability of being in the high state is equal to $d_h/(d_h + d_l)$. Second, we assume that every buyer receives an additional, idiosyncratic signal regarding the state of the world. Since the only purpose of these private signals is to order buyers' starting beliefs, we take them to be continuously distributed (which implies that the distribution of starting beliefs is continuous as well). In addition, we assume the signals to be sufficiently noisy that all buyers' starting beliefs are "close" to $d_h/(d_h + d_l)$. The questions of (a) how small precisely the range of starting beliefs needs to be and (b) why we do not want it to be larger will be answered in due course.

Why do buyers' beliefs about the state of the world (which we are also going to refer to as their respective *types*) matter at all? Given that the mass of new-born buyers is different in the two states, the same will (as we shall see) hold true for the degree of competition among buyers, i.e. winning an auction at a given price will be more valuable in the high state than in the low one. Hence, buyers would ideally like to condition their bids on d. As nobody of them actually knows d, θ is used as a proxy.

Consider next the evolution of buyers' beliefs over time: We assume that no information regarding the number of competing bidders or their bids is revealed to the losers of an auction. Nevertheless, they learn something about the state of the world: The fact that at least one competing bid must have been higher than their own one allows them to draw some conclusions about the state of the world. More specifically, let $G_k^t(b)$ denote the probability of winning an auction with bid b in state k and period t. This probability depends on the size and number of other buyers' bids. Assuming that $G_k^t(b)$ is known (for all b, t, and k), a bidder can, upon losing an auction in period $t-1$ with bid b, update her belief θ according to Bayes' rule:

$$z^t(\theta, b) = \frac{\theta \cdot (1 - G_h^{t-1}(b))}{\theta \cdot (1 - G_h^{t-1}(b)) + (1-\theta) \cdot (1 - G_l^{t-1}(b))}. \tag{1.1}$$

It bears emphasis that the current belief of a buyer contains all information she has ever received. For this reason, the belief of a buyer in period $t-1$ does not matter anymore once she has used it to compute that in period t.

Every buyer's goal is to maximize her expected lifetime-payoff which is equal to the difference between valuation and payment if she manages to win an auction and zero otherwise. Matters are complicated, however, by the fact that buyers can successively participate in more than one auction. What keeps the buyers' optimization problem tractable is our assumption that the matching of buyers to auctions occurs randomly and independently. As there are continuums of agents this assumption implies that the probability of any two buyers meeting more than once is zero. Therefore, it appears reasonable to posit that all buyers take $G_h^t(b)$ and $G_l^t(b)$ as given and never try to manipulate other buyers' beliefs.

As buyers' past beliefs do not contain any information over and above their current ones, all they need to know in order to bid optimally in period t is their belief in that period and the winning probability functions $G_h^t(b)$ and $G_l^t(b)$ in all future periods. The number of competing bidders is assumed to be unobservable and, hence, cannot be conditioned on. The recursive structure of the buyers' problem allows us to describe the value of being able to participate in the game by the Bellman equation

$$V^t(v, \theta) = \max_b U(b, v, \theta), \tag{1.2}$$

where the utility function $U(\cdot, \cdot, \cdot)$ is the same for all buyers and can be written as

$$U(b, v, \theta) = \int_0^b (v - x) g^t(x, \theta) dx + \delta \cdot L^t(b, \theta) \cdot V^{t+1}(z^{t+1}(\theta, b), v). \tag{1.3}$$

Here, $g^t(x, \theta) \equiv \theta \cdot g_h^t(x) + (1 - \theta) \cdot g_l^t(x)$ is the density function of winning bids as perceived by a buyer with belief θ and $L^t(b, \theta) \equiv \theta \cdot (1 - G_h^t(b)) + (1 - \theta) \cdot (1 - G_l^t(b))$ denotes the probability a buyer with belief θ attaches to not winning an auction with bid b. For now we simply conjecture $g^t(x, \theta)$ to exist, this claim will be confirmed in Proposition 1. The integration variable x denotes the highest competing bid. As apparent from equation (1.3), the utility function consists of two parts: The first one captures the utility drawn from the opportunity to win the next auction. The second one reflects the fact that bidders may, if they are lucky, get more than one chance to buy the good. Thus, every buyer has to trade off the risk of being eliminated (a good reason not to bid too low) against the opportunity cost of making a better deal in the future (a good reason not to bid too high).

In order to solve the buyers' problem we first have to find out more about the properties of $G_h^t(\cdot)$, $G_l^t(\cdot)$, and $V^t(\cdot, \cdot)$. This is the job of the next section.

1.2.4 Stocks and Flows of Buyers and Sellers

Our first goal in this subsection is to examine how the stocks of buyers and sellers evolve over time. Given our assumption that buyers are independently and randomly matched to sellers, the number of bidders participating in any given auction is a random variable following a Poisson distribution with parameter μ, where $\mu \equiv D/S$ denotes the ratio of buyers to sellers in the market.[9] The share of auctions not even

[9]Strictly speaking, the number of buyers per seller follows a Binomial distribution. As is well known, however, the Binomial distribution converges to the Poisson distribution as n (the number

attracting one single buyer is thus given by $e^{-\mu}$, the value of the Poisson distribution function at zero. Recall that the inflow of new buyers is equal to d_k, whereas the outflow consists of the winning bidders and a fixed fraction of the losers. The mass of buyers thus obeys the law of motion

$$D_k^t = d_k + \delta \cdot \left(D_k^{t-1} - S_k^{t-1} \cdot \left(1 - \exp(-S_k^{t-1}/D_k^{t-1})\right)\right) \tag{1.4}$$

The first term on the right hand side of equation (1.4) captures the mass of agents who get born in period t, the second one describes the agents who get inherited from the previous period. As for the sellers, we can write down the analogous equation

$$S_k^t = 1 + \delta \cdot \left(S_k^{t-1} - S_k^{t-1} \cdot \left(1 - \exp(-S_k^{t-1}/D_k^{t-1})\right)\right) \tag{1.5}$$

The assumption that buyers are independently matched to sellers entails a key advantage: It implies that $G_h^t(b)$ and $G_l^t(b)$ can be interpreted both as the distributions of highest competing bids *and* as the distributions of *winning* bids.

We have stressed above that all buyers take $G_h^t(b)$ and $G_l^t(b)$ as given. Nevertheless, both of them are endogenous, and our next goal will be to show how they derive from the model's fundamentals. In the general case of heterogeneous valuations, this is an extremely ambitious quest, because the agents already differ along another dimension, their belief about the state of the world. This makes it difficult to say anything about what their bidding function will look like.[10] With homogeneous valuations, on the other hand, it is natural to guess that, in equilibrium, higher types will submit higher bids. Why does this presumption appear to be reasonable? Higher beliefs should translate into higher bids, because agents who are relatively certain to be in the high state cannot expect to make a good bargain in the future, so the opportunity cost of winning is relatively low for them. In what follows we therefore assume that valuations are homogeneous and restrict attention to equilibria in strictly increasing strategies. Consequently, v is dropped as an argument from both $V(\cdot)$ and $U(\cdot, \cdot)$.

---

of buyers) goes to infinity while np (the number of buyers per seller) is bounded.

[10]The difficulties associated with the analysis of two-dimensional auctions are discussed in Pesendorfer and Swinkels (2000).

This restriction on the set of candidate equilibria simplifies the derivation of $G_h^t(b)$ and $G_l^t(b)$ considerably, as it implies that a buyer's bid always reveals her respective type. Thus, the question of how an agent updates her belief after having lost an auction with a certain (equilibrium) *bid* essentially reduces to how the agent updates her belief after having lost an auction as a certain *type*. Accordingly, we can separate the dynamics of buyers' types from their equilibrium bidding strategy, i.e. we do not have to solve the buyers' problem in order to derive the distribution function of their types in any given period. Formally, equation (1.1) can be written as

$$\zeta_t(\theta) = \frac{\theta \cdot (1 - \Gamma_h^{t-1}(\theta))}{\theta \cdot (1 - \Gamma_h^{t-1}(\theta)) + (1-\theta) \cdot (1 - \Gamma_l^{t-1}(\theta))} \tag{1.6}$$

where $\Gamma_k^t(\theta) \equiv G_k^t(b_t^*(\theta))$ denotes the distribution function of winning *types* (rather than bids), and $b_t^*(\cdot) : [0,1] \to \mathbb{R}_+$ is the buyers' equilibrium strategy in period t. Note that - given our focus on strictly increasing strategies - the probability of an agent with type $\hat{\theta}$ to win a given auction solely depends on (a) the mass of agents whose type is larger than $\hat{\theta}$ and (b) the mass of sellers. Formally,

$$\Gamma_k^t(\theta) = \exp(-\mu_k^t(\theta)) \equiv \exp(-D_k^t(\theta)/S_k^t), \tag{1.7}$$

where $D(\theta)$ denotes the mass of buyers whose type is larger than or equal to θ. This insight allows us to rewrite equation (1.6) further as

$$\zeta_t(\theta) = \frac{\theta \cdot (1 - \exp(-\mu_h^{t-1}(\theta)))}{\theta \cdot (1 - \exp(-\mu_h^{t-1}(\theta))) + (1-\theta) \cdot (1 - \exp(-\mu_l^{t-1}(\theta)))} \tag{1.8}$$

Before we can proceed, we need to impose one additional restriction on the model's equilibrium objects: We conjecture that in all periods an agent's posterior is a differentiable and strictly increasing function of her prior, i.e. $\zeta_t'(\theta) > 0 \; \forall \; \theta$. Thus, $\zeta_t(\cdot)$ is ensured to be invertible.

Our first step towards deriving the distribution of winning bids in any given period is to give a recursive formulation of the density function of *all* buyers' types (not only the winning ones) in period t, denoted as $\phi_t(\cdot)$. To this end, note first that our focus on strictly increasing strategies implies that $\phi_t(\cdot)$ exists if and only if this is the case for $g_t(\cdot)$. As mentioned above, we will show the distribution of buyers' types / bids to be differentiable in Proposition 1. The stock of buyers with belief θ in period t is made up of two groups: The buyers whose belief in period $t-1$ was $\zeta_t^{-1}(\theta) = z_t^{-1}(\theta, b_t^*(\theta))$ and who survived the end of that period and the buyers who are born in period t with a belief of θ. Leaving this latter group aside the mass of buyers with type smaller than of equal to θ is given by

$$D_k^t \cdot \Phi_k^t(\theta) \equiv D_k^{t-1} \cdot \delta \cdot \int_0^{\zeta_{t-1}^{-1}(\theta)} \phi_k^{t-1}(x) \cdot \left(1 - \Gamma_k^{t-1}\left(\zeta_{t-1}^{-1}(x)\right)\right) dx, \tag{1.9}$$

where $\Phi(\cdot)$ denotes the distribution function of all buyers' types (not just the winning ones).[11] Taking the derivative of the identity (1.9) we get that

$$D_k^t \cdot \phi_k^t(\theta) = \frac{D_k^{t-1} \cdot \phi_k^{t-1}\left(\zeta_{t-1}^{-1}(\theta)\right) \cdot \delta \cdot \left(1 - \Gamma_k^{t-1}\left(\zeta_{t-1}^{-1}(\theta)\right)\right)}{\zeta_{t-1}'(\zeta_{t-1}^{-1}(\theta))} \tag{1.10}$$

The denominator of the second term on the right hand side of (1.10) corrects for the fact that the difference between two similar (but not identical) priors may be bigger or smaller than the difference between the corresponding posteriors. One and the same mass of buyers can thus get either concentrated in a small or spread out over a large interval. The derivative in the denominator of (1.10) adjusts the density function by the ratio of the two differences.

Deriving the density function of *winning* types, $\gamma(\theta)$, is somewhat tricky, because the number of buyers participating in any given auction is a random variable. Letting n denote the number of bidders in an auction, $\gamma(\theta)$ can be written as

$$\gamma_t^k(\theta) = \sum_{n=1}^{\infty} \exp(-\mu_k^t) \cdot \left(\mu_k^t\right)^n / n! \cdot n \cdot \left[\Phi_t^k(\theta)\right]^{n-1} \cdot \phi_t^k(\theta), \tag{1.11}$$

We now have collected all ingredients which are necessary to formally define an equilibrium of our model. This task will be tackled in the next section.

[11] In our notation Greek letters are always associated with buyers' *types*, while Latin letters are associated with their *bids*.

1.3 Steady State Equilibrium

1.3.1 Steady State Type Distributions

Our main goal in this section is to prove equilibrium existence and - within a certain class of equilibria - uniqueness for the model presented in the previous section. In addition, we want to characterize some properties of this equilibrium. For the sake of tractability our analysis will focus on *steady-state equilibria*, i.e. on equilibria in which the model's endogenous variables remain constant over time. Accordingly, we will sometimes talk about the "steady state version" (*ssv*) of an equation, meaning a version of the respective equation where all time indices have been dropped. Before we can start proving the existence of a steady state equilibrium we first have to make precise what we are talking about. This is the job of the following

Definition 1 *A steady state equilibrium of the model is a tuple* $\{D_k,\, S_k,\, \phi_k(\theta),\, \gamma_k(\theta),\, b^*(\theta),\, f_k(b),\, g_k(b),\, V(\theta);\, k = h, l\}$ *such that*

- D_h, D_l, S_h, *and* S_l *satisfy the ssv's of equations (1.4), (1.5) in both states.*
- $\phi_h(\cdot)$, $\phi_l(\cdot)$, $\gamma_h(\cdot)$, *and* $\gamma_l(\cdot)$ *satisfy the ssv's of equations (1.6), (1.10), and (1.11).*
- $b^*(\theta)$ *is payoff maximizing, i.e.* $b^*(\theta) \in \arg\max U(b, \theta)\ \forall\ \theta$
- $g_k\left(b^{*\prime}(\theta)\right) = \gamma_k(\theta)$ *and* $f_k\left(b^{*\prime}(\theta)\right) = \phi_k(\theta)$, $k = h, l$
- $V(\theta)$ *satisfies the ssv of equation (1.2)*

As mentioned before, we will restrict attention to equilibria in strictly increasing strategies. The plan of this section is as follows: We first — in Lemma 1 and Propositions 1 to 3 — examine the mass of buyers and the distribution of their types. We then — in Propositions 4 to 6 — turn to buyers' optimal bidding behavior and finally — in Proposition 7 — analyze the value function $V(\theta)$. Thus, we once again exploit the fact that under the assumption of strictly increasing strategies we do not need to know buyers' equilibrium bidding function in order to characterize the updating process.

In a steady state the inflow of new buyers needs to equal the outflow of exiting ones (in both states of the world), i.e.

$$d_k - (1-\delta) \cdot D_k - S_k \cdot \delta \cdot (1 - \exp(-D_k/S_k)) = 0 \tag{1.12}$$

For every mass of buyers $D_k \in [d_k, d_k/(1-\delta)]$ we can find a unique value of S_k such that equation (1.12) holds. Let $P_k : (d_k, d_k/(1-\delta)] \to \mathbb{R}_+$ denote the function which gives us this value. The mass of sellers is in steady state if and only if

$$1 - (1-\delta) \cdot S_k - S_k \cdot \delta \cdot (1 - \exp(-D_k/S_k)) = 0 \tag{1.13}$$

Let $Q : (d_k, d_k/(1-\delta)] \to \mathbb{R}_+$ denote the function which, for any given mass of buyers, gives us the value of S which ensures that the mass of sellers is in steady state. We can now state our first Lemma:

Lemma 1: *There is a unique tuple (D_k, S_k) such that equations (1.12) and (1.13) both hold, i.e. a unique steady state with respect to the masses of buyers and sellers exists. Moreover, $\mu_h > \mu_l$.*

Proof. The two sister functions $P_k(\cdot)$ and $Q(\cdot)$ exhibit the following set useful properties: First, $\lim_{D \to d_k} P_k(D) = \infty$, whereas $\lim_{D \to d_k} Q(D) < \infty$. Second, $P_k(d_k/(1-\delta)) = 0$, whereas $Q(d_k/(1-\delta)) > 0$. Third, $P'_k(D) < Q'(D) \; \forall \; D$. The first two properties imply that - by the intermediate value theorem (de la Fuente, 2000, p.76) - there exists an equilibrium. The third property tells us that this equilibrium must be unique. In addition, we know that $P_h(D) > P_l(D)$, implying that $D_h > D_l$, whereas $S_h < S_l$. It immediately follows that $\mu_h > \mu_l$, i.e. the ratio of buyers to sellers is larger in the high than in the low state. ∎

Both the steady state's uniqueness and the effect of a change in d_k are graphically depicted in Figure 1.

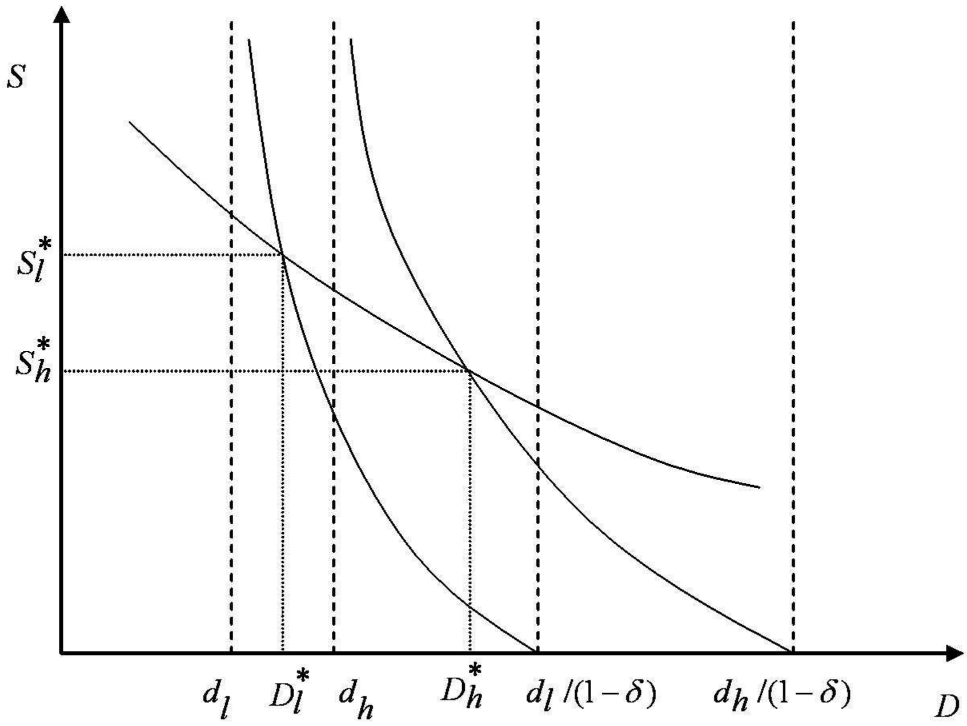

Figure 1: Equilibrium Loci with Respect to Buyers and Sellers

Deriving the steady state density functions of buyers' types is a much more ambitious quest, and we cannot complete it without confining the space of candidate functions. In what follows we will therefore restrict our attention to equilibria in which the function $1-\Gamma(\theta)$ has the *Monotone Likelihood Ratio Property* (MLRP), i.e. in which the derivative of $(1-\Gamma_h(\theta))\,/\,(1-\Gamma_l(\theta))$ with respect to θ is strictly positive. This restriction has two important implications: First, it implies that an agent's posterior is a strictly increasing function of her prior, i.e. no agents' types can ever "cross". (We have already conjectured this property of $\zeta(\theta)$ in our derivation of equation (1.9)). Second, the agents always update their beliefs upwards. Why does this need to be the case? We have just seen that $\mu_h > \mu_l$, implying that $(1-\Gamma_h(0))\,/\,(1-\Gamma_l(0)) > 1$. By the MLRP this allows us to conclude that $(1-\Gamma_h(\theta))\,/\,(1-\Gamma_l(\theta)) > 1\ \forall\ \theta$. Hence, agents always update their beliefs upwards. Let the highest and lowest type within the first generation of buyers be denoted as $\underline{\theta}$ and $\overline{\theta}$, respectively. For $\overline{\theta}-\underline{\theta}$ small enough it must be true that $\zeta(\overline{\theta}) > \underline{\theta}$.[12] The distribution functions of buyers' types

[12]The assumption that $\zeta(\overline{\theta}) > \underline{\theta}$ is not essential, but it simplifies the exposition of the model.

then looks as follows: There is perfect sorting with respect to agents' age, i.e. agents who have spent more time in the market attach a higher probability to being in the high state. Moreover, there can (and generally will) be gaps in between any two generations, i.e. there are intervals of beliefs which are "leapfrogged" by all buyers so that nobody holds them. We are now in a position to derive the equilibrium functions $\phi_l(\theta)$ and $\phi_h(\theta)$.

Intuitively, our strategy looks as follows: From Lemma 1 we know $\mu_h(\theta)$ and $\mu_l(\theta)$ for all $\overline{\theta} > \theta > \underline{\theta}$. Since both the survival probability of an agent and her posterior type solely depend on $\mu_h(\theta)$ and $\mu_l(\theta)$, our knowledge about the first generation buyers (and the total number of buyers per auction) is sufficient to determine the distribution of the *second generation* buyers. Since no agent ever updates her belief downwards, this process can be repeated over and over again until we know the entire distribution of buyers' types.

Proposition 1: *Under the assumptions that (a) the agents use strictly increasing strategies, (b)* $1 - \Gamma(\theta)$ *has the MLRP, and (c)* $\zeta(\overline{\theta}) > \underline{\theta}$*, there exists - for given* S_k and D_k *- a unique pair of functions* $\phi_l(\theta)$ *and* $\phi_h(\theta)$ *which is consistent with our definition of an equilibrium.*

Proof. The proof is by construction: Consider first the sequence $\{\zeta^n(\theta)\}_{n=0}^{\infty}$. The nth element of this sequence gives us the type of an agent who, starting from a belief of θ, has lost n auctions. Accordingly, $\zeta^{n+1}(\theta)$ is recursively defined as

$$\zeta^{n+1}(\theta) = \zeta(\zeta^n(\theta)) \tag{1.14}$$

where $\zeta^0(\theta)$ denotes the (known) belief of a first-generation buyer and $\zeta(\theta)$ is the steady state analogue of $\zeta_t(\theta)$ (as defined in equation (1.8)), so $\zeta^{n+1}(\theta)$ solely depends on $\zeta^n(\theta)$ and on $\mu_h(\zeta^n(\theta))$ and $\mu_l(\zeta^n(\theta))$. From equation (1.4) we know that

$$\mu_k\left(\zeta^{n+1}(\theta)\right) = \delta \cdot (\mu_k(\zeta^n(\theta)) - 1 + \exp(\mu_k\left(\zeta^n(\theta)\right)) \tag{1.15}$$

Hence, $\mu_k(\zeta^{n+1}(\theta))$ only depends on $\mu_k(\zeta^n(\theta))$. Therefore, once $\mu_h(\zeta^n(\theta))$, $\mu_l(\zeta^n(\theta))$, and $\zeta^n(\theta)$ are known, we can compute $\zeta^{n+1}(\theta)$, $\mu_h\left(\zeta^{n+1}(\theta)\right)$ and $\mu_l\left(\zeta^{n+1}(\theta)\right)$. This allows us to iterate on equations (1.14) and (1.15).

If there exists no n such that $\zeta^n(\underline{\theta}) < \theta_0 < \zeta^n(\overline{\theta})$, we know that $\phi_l(\theta_0) = \phi_h(\theta_0) = 0$. If, on the other hand, there is some n such that $\zeta^n(\underline{\theta}) < \theta_0 < \zeta^n(\overline{\theta})$ (denoted as $\bar{n}$), we have to look for $\zeta^{-\bar{n}}(\theta_0)$, i.e. the "ancestor" of θ_0. As we know $\zeta^{-\bar{n}}(\theta_0)$ to lie somewhere between $\underline{\theta}$ and $\overline{\theta}$ and $\zeta^n(\theta)$ is a continuous and strictly increasing function of θ, $\zeta^{-\bar{n}}(\theta_0)$ is known to exist and to be unique.

Once we know $\zeta^{-\bar{n}}(x_0)$, $\phi_l(\theta_0)$ and $\phi_h(\theta_0)$ can be found by considering the sequence $\{\phi_k(\zeta^n(\theta))\}_{n=1}^{\infty}$, which is - as we know from equation (1.10) - recursively defined as

$$\phi_k(\zeta^n(\theta)) = \frac{\phi_k(\zeta^{n-1}(\theta)) \cdot \delta \cdot \left(1 - \Gamma(\zeta^{n-1}(\theta))\right)}{\zeta'\left(\zeta^{n-1}(\theta)\right)}, \tag{1.16}$$

where $\zeta^{-\bar{n}}(\theta_0)$ is used as the starting value. The $\bar{n}$th element of this sequence gives us the desired densities. Note that, as $(1 - \Gamma_h(0)) / (1 - \Gamma_l(0)) > 1$, $\zeta'(\theta)$ is bounded away from zero for all $\theta < 1$. Hence, the finiteness of $\phi_k(\cdot)$ is preserved by iterations on equation (1.16). Therefore, it is admissible to characterize the distribution of buyers' types by a density function (as we have conjectured above). ■

The basic intuition behind Proposition 1 is that, since any loss provides evidence in favor of the high state, no agent ever updates her belief downwards. Thus, we can construct the distribution function of buyers' types "from bottom to top".

Having derived the equilibrium distributions of buyers' types our next goal is to say more about their properties. In particular, we want to check whether the distribution function of winning types generated by $\phi_l(\theta)$ and $\phi_h(\theta)$ has the MLRP (as we have assumed above). To this end, note first that $\phi_l(\theta)$ and $\phi_h(\theta)$ themselves necessarily have the MLRP. Why does this need to be the case? Every buyer's belief about the state of the world results from a sequence of signals, it depends on the likelihood of that sequence occurring in each of the two states. If a sequence is R-times as likely to occur in the high state than in the low one, all buyers who have experienced it will hold a belief of $\theta = R/(R + 1)$. At the same time, we know that the mass of these buyers must be R-times as large in the high state than in the low one. Hence, θ, $\phi_h(\theta)$, and $\phi_l(\theta)$ are linked by the following equation: $(D_h \cdot \phi_h(\theta)) / (D_l \cdot \phi_l(\theta)) = \theta/(1 - \theta)$. Since the right hand side of this equation is increasing in θ, $\phi(\theta)$ must have the MLRP. Our next goal is to show that the same holds true for $\gamma(\theta)$.

Proposition 2: *$\gamma(\theta)$ has the MLRP if this is the case for $\phi(\theta)$.*[13]

Proof. Note first that, as the number of buyers in any given auction is Poisson-distributed, the distribution function of the first order statistic of buyers' types in state k is given by

$$\Gamma_k(\theta) = \sum_{n=0}^{\infty} \frac{e^{-\mu_k} \cdot \mu_k^n}{n!} \cdot [\Phi_k(\theta)]^n \tag{1.17}$$

Taking the derivative of this expression we get that

$$\gamma_k(\theta) = \sum_{n=1}^{\infty} \frac{e^{-\mu_k} \cdot \mu_k^n}{(n-1)!} \cdot [\Phi_k(\theta)]^{n-1} \cdot \phi_k(\theta) \tag{1.18}$$

In order for $\gamma(\theta)$ to have the MLRP it must hold that $\gamma_h'(\theta)/\gamma_h(\theta) > \gamma_l'(\theta)/\gamma_l(\theta)$. As for $\gamma_k'(\theta)$, we know that

$$\gamma_k'(\theta) = \sum_{n=1}^{\infty} \frac{e^{-\mu_k} \cdot \mu_k^n}{(n-1)!} [\Phi_k(\theta)]^{n-1} \cdot \phi_k'(\theta) + \sum_{n=2}^{\infty} \frac{e^{-\mu_k} \cdot \mu_k^n}{(n-2)!} [\Phi_k(\theta)]^{n-2} \cdot [\phi_k(\theta)]^2$$

The second term on the right hand side of this equation can be written as

$$\mu_k \cdot \phi_k(\theta) \cdot \sum_{n=2}^{\infty} \frac{e^{-\mu_k} \cdot \mu_k^{n-1}}{(n-2)!} \Phi_k^{n-2}(\theta) \cdot \phi_k(\theta) \tag{1.19}$$

which is equal to $\mu_k \cdot \phi_k(\theta) \cdot \gamma_k(\theta)$. Thus, a sufficient condition for $\gamma(\theta)$ to have the MLRP is that

$$\frac{\gamma_h'(\theta)}{\gamma_h(\theta)} = \mu_h \cdot \phi_h(\theta) + \frac{\phi_h'(\theta)}{\phi_h(\theta)} > \mu_l \cdot \phi_l(\theta) + \frac{\phi_l'(\theta)}{\phi_l(\theta)} = \frac{\gamma_l'(\theta)}{\gamma_l(\theta)} \tag{1.20}$$

A sufficient condition for this inequality to hold is that (a) $\phi_h'(\theta)/\phi_h(\theta) > \phi_l'(\theta)/\phi_l(\theta)$) and (b) $\mu_h \cdot \phi_h(\theta) > \mu_l \cdot \phi_l(\theta)$. Condition (a) clearly is satisfied as $\phi(\theta)$ has the MLRP. Condition (b) posits that the mass of buyers within any interval of types is larger in the high state than in the low one. From our discussion above we know that this condition is tantamount to requiring that no buyer holds a belief of less than one half. If all newborn buyers' types are sufficiently close to $d_h/(d_h + d_l)$, this is indeed the case. ■

[13]It is important to recall that some auctions do not even attract a single buyer. Therefore, the distribution of winning bids has a mass point at zero. In Lauermann and Merzyn (2006) we show that $\Gamma_l(0)/\Gamma_h(0) > \gamma_l(\theta)/\gamma_h(\theta)$ for some θ, i.e. even a generalized MLRP does not hold at $\theta = 0$. This however, does not affect the agents' bidding strategy, for if an agent is the only bidder, she will get the item for free anyway.

The final step is to show that the MLRP of $1 - \Gamma(\theta)$ results from that of $\gamma(\theta)$.

Proposition 3: *If $\gamma(\theta)$ has the MLRP, this is also true of $1 - \Gamma(\theta)$. Moreover, $(1-\Gamma_h(\theta))/(1-\Gamma_l(\theta)) > \gamma_h(\theta)/\gamma_l(\theta)$ whenever the latter ratio is defined (i.e. whenever $\gamma_l(\theta) > 0$).*

Proof. [14] The derivative of $(1 - \Gamma_h(\theta))/(1 - \Gamma_l(\theta))$ (whenever it exists) is given by

$$\frac{d}{d\theta}\frac{1-\Gamma_h(\theta)}{1-\Gamma_l(\theta)} = \frac{\gamma_l(\theta)\cdot(1-\Gamma_h(\theta)) - \gamma_h(\theta)\cdot(1-\Gamma_l(\theta))}{(1-\Gamma_l(\theta))^2} \tag{1.21}$$

This derivative is positive if and only if $(1 - \Gamma_h(\theta))/(1 - \Gamma_l(\theta) > \gamma_h(\theta)/\gamma_l(\theta)$. Thus we can prove both parts of our proposition at once.

$$\frac{1-\Gamma_h(\theta)}{1-\Gamma_l(\theta)} = \int_\theta^1 \frac{\gamma_h(x)\cdot\gamma_l(x)}{\gamma_l(x)\cdot(1-\Gamma_l(\theta))}dx > \int_\theta^1 \frac{\gamma_h(\theta)\cdot\gamma_l(x)}{\gamma_l(\theta)\cdot(1-\Gamma_l(\theta))} = \frac{\gamma_h(\theta)}{\gamma_l(\theta)} \tag{1.22}$$

This inequality completes our proof. Note that $\gamma_h(x)/\gamma_l(x) > \gamma_h(\theta)/\gamma_l(\theta)$ for all $x > \theta$, because $\gamma_h(\theta)/\gamma_l(\theta)$ is increasing in θ. ■

Propositions 2 and 3 do not only show that there exists a steady state distribution of buyers' beliefs such that $1 - \Gamma(\theta)$ has the MLRP, but also that in this steady state $\gamma(\theta)$ likewise has the MLRP. This result will prove useful in our analysis of buyers' optimal bidding strategies which is the subject of the next section.

1.3.2 Optimal Bidding Strategies and Equilibrium Payoffs

Our first goal in this section is to characterize the shape of the value function $V(\theta)$. As $\Gamma_h(\theta)$ first-order stochastically dominates $\Gamma_l(\theta)$ and the agents use strictly increasing strategies, $V(\theta)$ is decreasing, i.e. the more likely an agent is to be in the high state, the worse off she is. In addition, the following proposition reveals that $V(\theta)$ is convex.

Proposition 4: *The value function is convex, i.e. $\alpha \cdot V(\theta_1) + (1-\alpha)\cdot V(\theta_2) \geq V(\alpha\cdot\theta_1 + (1-\alpha)\cdot\theta_2)\ \forall\ \theta_1, \theta_2, \alpha \in [0,1]$*

Proof. Let $W_k(\theta)$ denote the expected lifetime payoff of a buyer who is known to be in state k, but uses the equilibrium strategy of type θ. The expected lifetime payoff of a buyer

[14] The proof is adopted from Milgrom (1981, p. 926)

who is of type θ_1, but behaves (i.e. bids) as if he was of type θ_2 (denoted as $W(\theta_2, \theta_1)$ can then be rewritten as

$$W(\theta_2, \theta_1) = \int_0^{b_2^*} (v - x) g(x, \theta_1) dx + \delta \cdot L(b_2^*, \theta_1) \cdot W\left(b^*\left(z(\theta_2, b_2^*)\right), z(\theta_1, b_2^*)\right) \tag{1.23}$$

where $b_2^* = b^*(\theta_2)$ and $W(\theta_2, \theta_1) = \theta_1 \cdot W_h(\theta_2) + (1 - \theta_1) \cdot W_l(\theta_2)$. Let us now get back to our initial problem: Defining $\theta_3 = \alpha \cdot \theta_1 + (1 - \alpha) \cdot \theta_2$ we have

$$\alpha \cdot W(\theta_3, \theta_1) + (1 - \alpha) \cdot W(\theta_3, \theta_2) = V(\theta_3) \tag{1.24}$$

In addition, it by definition holds true that $V(\theta_1) \geq W(\theta_3, \theta_1)$ and $V(\theta_2) \geq W(\theta_3, \theta_2)$ which completes the proof of the proposition. ■

The convexity of the value function mirrors the fact that the value of information in our game cannot be negative. Hence, whenever there is uncertainty about the true state of the world, well informed buyers have a comparative advantage over their competitors. Now we can finally turn to the problem of examining buyers' bidding behavior. There are two different ways of approaching this issue, and since they both have distinct advantages, we will pursue them subsequently.

One way to find the first-order condition for a bid to be optimal is to maximize the utility function (1.3) subject to the updating formula (1.1), i.e. to substitute for $z(\theta)$ in (1.3) and to then look for the optimum with respect to b. This approach is complicated by the fact that the value function may not be differentiable everywhere. Being a convex function, however, $V(\theta)$ does have a subderivative everywhere which we denote as $\partial V(\theta)$.[15] Thus, we can find a candidate equilibrium bidding function by taking the subderivative of the utility function (1.3) with respect to b and looking for the bid b^* at which it is zero. Letting $K \equiv \theta \cdot (1 - G_h(b))$, $L \equiv \theta \cdot (1 - G_h(b)) + (1 - \theta) \cdot (1 - G_l(b))$, $K' \equiv -\theta \cdot g_h(b)$, and $L' \equiv -\theta \cdot g_h(b) - (1 - \theta) \cdot g_l(b)$ we have that

$$\frac{d}{db} \int_0^b (x - v) \cdot (\theta \cdot g_h(b) + (1 - \theta) \cdot g_l(b))\, dx + \delta \cdot L \cdot V\left(K/L\right)$$

[15] The subderivative of the function $f : X \to \mathbb{R}$ at x_0, *denoted as* $\partial f(x_0)$, *is the set of all numbers* $p \in \mathbb{R}$ *such that* $f(x_0) + p \cdot (x - x_0) \leq f(x)$ *for all* $x \in X$ (de la Fuente, 2000, p. 248).

$$= (b - v) \cdot L' + \delta \cdot \left(L' \cdot V(K/L) + L \cdot V'(K/L) \cdot (K' \cdot L - L' \cdot K) / L^2 \right)$$

$$= L' \left(b - v + \delta \cdot V(K/L) + V'(K/L) \cdot (K'/L' - K/L) \right) = 0 \qquad (1.25)$$

The resulting first-order condition can be written as

$$b^*(\theta) = v - \delta \cdot \left(V\left(\zeta(\theta)\right) + \partial V\left(\zeta(\theta)\right) \cdot \left(\lambda(\theta) - \zeta(\theta)\right) \right) \qquad (1.26)$$

where $\lambda(\theta) = \theta \cdot \gamma_h(\theta) \cdot \left(\theta \cdot \gamma_h(\theta) + (1-\theta) \cdot \gamma_l(\theta)\right)^{-1}$ is the posterior of an agent with prior θ who has observed that the highest competing type in an auction was equal to her own one. The term in brackets on the right hand side of equation (1.26) is graphically depicted in Figure 2.

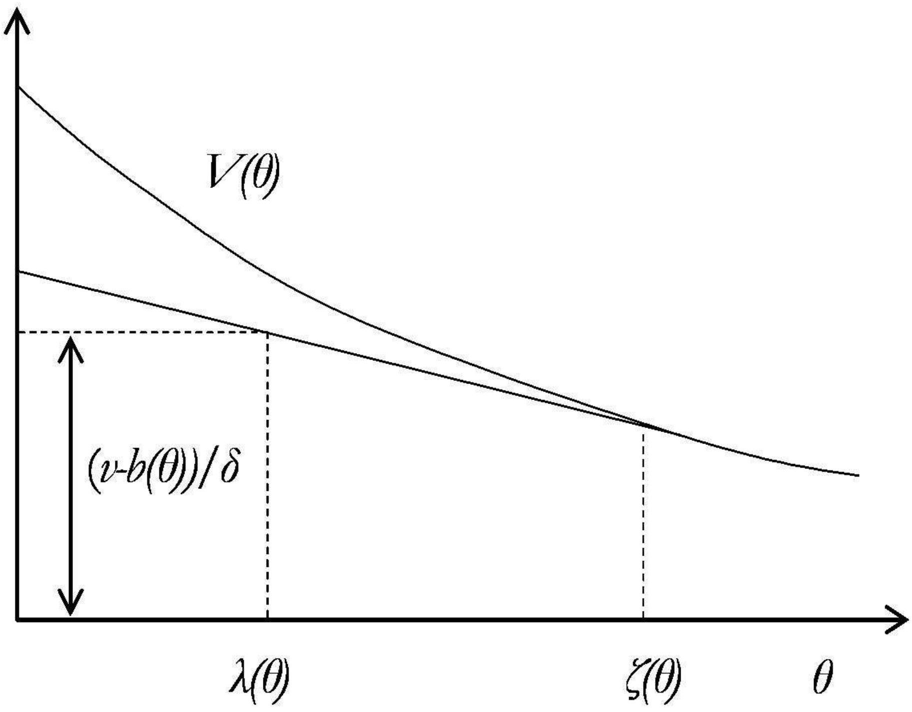

Figure 2: The Relationship between $V(\theta)$ and $b^*(\theta)$

Equation (1.26), however, is just a first-order condition, and, therefore, the strategy it prescribes is just an equilibrium *candidate.* We still need to verify that this candidate indeed constitutes an equilibrium. This is what the proof of our next proposition does.

Proposition 5: *Symmetric equilibrium bidding strategies are given by $b^*(\theta)$, i.e. if all buyers bid according to $b^*(\theta)$, the value function $V(\theta)$ is maximized by that choice.*

Proof. What we want to prove here is that, whenever all buyers bid according to $b^*(\theta)$, the subderivative of a buyer's payoffs with respect to b is positive for all $b < b^*(\theta)$, whereas it is negative for all $b > b^*(\theta)$. To this end, consider first two different beliefs θ_1 and θ_2, where $\theta_1 > \theta_2$. Note that

$$b^*(\theta_1) > v - \delta \cdot (V(\zeta(\theta_2)) + \partial V(\zeta(\theta_2)) \cdot (\lambda(\theta_1) - \zeta(\theta_2))) > b^*(\theta_2) \tag{1.27}$$

The first inequality sign in (1.27) follows from the convexity of $V(\theta)$ and the fact that $\xi(\theta)$ is - by Proposition 4 - smaller than $\zeta(\theta)$, the second one is simply due to $\lambda(\theta)$ being increasing in θ. Thus, the bidding strategy $b^*(\cdot)$ is strictly increasing in θ. This confirms our guess that, in the equilibrium we consider here, all buyers employ strictly increasing strategies. Let's now look at what happens when a buyer unilaterally deviates from the strategy given by (1.26) and bids as if her belief was θ_2, even though it is θ_1, where, again, $\theta_1 > \theta_2$. (Thus, she underbids, as $b^*(\theta)$ is increasing in θ.) The subderivative of this buyer's payoffs with respect to b is given by

$$v - b_2^* - \delta \cdot (V(z(\theta_1, b_2^*)) + \partial V(z(\theta_1, b_2^*)) \cdot (l(\theta_1, b_2^*) - z(\theta_1, b_2^*))) > 0 \tag{1.28}$$

Note that (by construction) the subderivative would be zero if the buyer's type was θ_2. The inequality sign follows from the fact that $z(\theta_1, b_2^*) > \zeta(\theta_2)$ and $l(\theta_1, b_2^*) > \lambda(\theta_2)$. The argument that a buyer who unilaterally over-bids will gain from an incremental decrease of her bid is analogous. ■

As we have hinted at above, there is an alternative way of deriving buyers' equilibrium bidding strategies. It is based on the function $W(\cdot, \cdot)$ which we have introduced in the proof of Proposition 4. (Recall that $W(\theta_2, \theta_1)$ is the expected lifetime utility of a buyer who is of type θ_1, but bids as if she was of type θ_2.)

Proposition 6: The buyers' equilibrium bidding function can be written as $b^*(\theta) = v - \delta \cdot W(\zeta(\theta), \lambda(\theta))$

Proof. We can express the utility function (1.3) as follows:

$$U(b,\theta)=\int_0^b (v-x)g(x,\theta)dx+\delta\cdot\int_b^v W(z(\theta,b),l(\theta,x))g(x,\theta)dx \tag{1.29}$$

Equation (1.29) differs from its counterpart (1.3) in that we consider all highest competing bids between b and v *separately.* This approach may appear somewhat unreasonable, because the buyers only observe that the highest competing bid was higher than their own one. As will become clear very soon, it nevertheless makes sense for them to distinguish between losses against different types. Taking the derivative of $U(\cdot,\cdot)$ (as defined in (1.29)) with respect to b gives us the following first-order condition:

$$b-v+\delta\cdot W(z(\theta,b),l(\theta,b))=\frac{d}{db}\int_b^v W(z(\theta,b),l(\theta,x))g(x,\theta)dx \tag{1.30}$$

The term on the right hand side of (1.30) is equal to

$$W_1(z(\theta,b),z(\theta,b))\cdot\frac{d}{db}z(\theta,b)\cdot(1-G(x,\theta)), \tag{1.31}$$

$W(\cdot,\cdot)$ does not need to be differentiable in its first argument everywhere. It will, however, be differentiable at $W(\theta,\theta)$ (in both arguments) whenever $V(\theta)$ is differentiable at θ. Being a monotone function, $V(\theta)$ is known to be differentiable *almost* everywhere (Rockafellar, 1970, p. 246). This does not only ensure that the derivative in (1.30) exists. From the envelope theorem of Milgrom and Segal (2002) we also know that $W_1(\theta_2,\theta_1|\theta_2=\theta_1)=V(\theta_1)-W_2(\theta_2,\theta_1|\theta_2=\theta_1)=0$, i.e. a buyer's expected lifetime payoff is unaffected by marginal deviations from her optimal strategy. Thus, the bidding strategy prescribed by the first-order condition (1.30) simplifies to

$$b^*(\theta)=v-\delta\cdot W(\zeta(\theta),\lambda(\theta)) \tag{1.32}$$

Once again, this first-order condition is insufficient for $b^*(\cdot)$ to describe an equilibrium-strategy. However, rather than checking that it is indeed an equilibrium for all buyers to bid according to (1.32) - which would be an exercise similar to that performed in the proof of Proposition 6 -, we will instead show that the two formulations given in (1.26) and (1.32) are, in fact, equivalent. To this end, note that $W(\theta_2,\theta_1)$ is linear in θ_1. Since, in addition, we know from the envelope theorem that $\partial V(\theta)=\partial_2 W(\theta_2,\theta_1|\theta_2=\theta_1)$, it must be true that

$$W(\zeta(\theta),\lambda(\theta))=V(\zeta(\theta))+\partial V(\zeta(\theta))\cdot(\lambda(\theta)-\zeta(\theta)) \tag{1.33}$$

This proves that (1.26) and (1.32) describe the same strategies and, hence, completes the proof of the proposition. ■

The formulation in (1.32) has two advantages over that in (1.26). First, it reveals the analogy between the auctions considered in our model and standard common value auction: In both cases the value of winning the item for sale depends not only on the winner's type, but also on the types of her opponents. The agents account for this interdependency by equating their bids to the expected value of winning the auction conditional on the highest competing type being precisely as high as their own one. To put it in the words of Krishna (2002):

> *A bidder with signal θ is asked to bid an amount $b(\theta)$ such that if he were to just win the auction with that bid - if the highest competing bid, and hence the price, were also $b(\theta)$ - he would just "break even".*

In the standard common value model this strategy does not appear peculiar; in our model, however, it requires the agents to condition their bids on a case of which they may (in case of a loss) never know whether it has occurred or not. This leads them to consider situations in which their perception of how likely they are to be in the high state of the world differs from the actual probability of being there. Nevertheless, their reasoning is exactly the same as that of bidders in a standard CV auction.

The second advantage of the formulation in (1.32) is that it will prove helpful in establishing the existence of a unique steady state equilibrium. What we need to show is the following:

Proposition 7: *There exists a unique function $W(\theta_1, \theta_2)$ which fulfills the Bellman equation (1.29) and is consistent with the equilibrium bidding function (1.32).*

Proof. The Bellman equation (1.29) can be viewed as an operator (denoted as T) which maps value functions into value functions. Letting W_1 and W_2 denote two different value functions, we want to show that T has the following two properties:

1. If $W_1(\theta_1, \theta_2) \leq W_2(\theta_1, \theta_2) \; \forall \; \theta_1, \theta_2$, then $T\left[W_1(\theta_1, \theta_2)\right] \leq T\left[W_2(\theta_1, \theta_2)\right] \; \forall \; \theta_1, \theta_2$

2. $T\left[W(\theta_1, \theta_2) + \alpha\right] \leq T\left[W(\theta_1, \theta_2)\right] + \alpha$

Ad 1 ("Monotonicity)": If $W_1(\theta_1, \theta_2) \leq W_2(\theta_1, \theta_2)\ \forall\ \theta_1, \theta_2$, then - by the bidding function 1.32 - we must have that $b_1^*(\theta) \geq b_2^*(\theta)\ \forall\ \theta$. Thus, no matter whether a buyer wins or loses an auction, she always benefits from low values of $W(\cdot\cdot)$. Ad 2 ("Discounting"): Note that the derivative of $b^*(\cdot)$ with respect to the value function is $-\delta$. Since the derivative of T with respect to $W(z(\theta, b), l(\theta, x))$ conditional on a loss is δ, while that with respect to $b^*(\cdot)$ conditional on a win is equal to minus one, a one-point increase of $W(\theta_1, \theta_2)\ \forall\ \theta_1, \theta_2$ leads to an increase of $T[\cdot]$ by $\delta < 1$. Taken together, the two properties of *Monotonicity* and *Discounting* mean that T satisfy Blackwell's sufficient conditions for a contraction (de la Fuente, 2000). It then follows from the contraction mapping theorem that the Bellman equation has a unique solution. ■

Since there is a one-to-one mapping from $W(\cdot, \cdot)$ to $V(\cdot)$, Proposition 7 also proves that equation (1.2) has a unique solution. Thus, it completes the proof of our claim that the model has a steady state equilibrium which is unique in the class of equilibria under consideration.

1.4 Sellers' Optimal Information Policy

In the previous section we have made the - rather extreme - assumption that a buyer who unsuccessfully participates in an auction does not learn anything about her competitors except for the fact that one of them bid higher than she did. Sellers were forced to stay silent and had no chance to provide buyers with any additional information. In this section we are going to relax this assumption by letting sellers choose between two different informational regimes: They may either conduct their auctions in an intransparent way (as in the previous section) or, alternatively, reveal the winning bid after each auction to the respective participants.[16] The question we are going to ask is the following one: If sellers are independent of each other and behave as revenue-maximizers, which regime will they choose? What makes this ques-

[16] Note that, since we consider second-price auctions, the winning bid is not identical to the price. Assuming that the latter is revealed would lead to asymmetric learning on the part of the losing bidders, because one of them would just learn his own bid.

tion relatively simple is the fact that there is a continuum of sellers and matches occur independently and randomly. Hence, sellers will take the functions $\gamma_h(\theta)$ and $\gamma_l(\theta)$ as given. It thus appears - at least for now - reasonable to do the same and to directly proceed with an analysis of buyers' optimal bidding behavior.

If the winning bid is announced after an auction (to the respective participants) buyers' utility function looks as follows:

$$U(b,\theta) = \int_0^b (v-x)g(x,\theta)dx + \delta \cdot \int_b^v V(l(\theta,x))g(x,\theta)dx \tag{1.34}$$

Taking the derivative of this function with respect to b, setting it to zero and solving for b gives us the buyers' optimal bidding strategy under the transparent regime:

$$b^*(\theta) = v - \delta \cdot V(l(\theta, b^*(\theta))) \tag{1.35}$$

The fact that it is indeed optimal for the buyers to follow the strategy prescribed by (1.35) follows - once again - from an argument analogous to that used in the proof of Proposition 5. We are now in a position to state the major proposition of the current section:

Proposition 8: *For given functions $\gamma_h(\theta)$ and $\gamma_l(\theta)$ buyers always bid (weakly) less under the transparent regime than they do under the intransparent regime. In other words, it never pays for a seller to reveal the winning bid after an auction.*

Proof. The proof is based on a simple comparison of the continuation values in the two regimes. In the transparent regime the losers of an auction get comparatively well informed about the state of the world. Thus, they are able to choose their bids in subsequent auctions more precisely. This means that their continuation payoff is higher than that of the buyers who have not learned the winning bid. As a result, they submit lower bids. Formally, $W(z(\theta, b^*(\theta)), l(\theta, b^*(\theta))) \leq V(l(\theta, b^*(\theta)))$. Since buyers always choose the bids which are optimal given their current beliefs, a buyer's payoff is maximized if her belief corresponds to the actual probability of being in the high state. ■

Proposition 8 might come as a surprise to those whose intuition - derived from the Linkage Principle - holds that the seller of common value good should always reveal as much information about that good as possible. The contrast between our result and the Linkage Principle can, however, easily be explained: In our model the buyers become only informed *after* the auction has taken place.[17] Thus, the only effect of announcing the winning bid is on the continuation value of the losing bidders. Since equilibrium bids depend negatively on that value, no seller will ever want to provide the losers of her auction with any information about the state of the world. By doing so, she would only make losing more attractive, thereby removing incentives to bid high.[18] Hence, it is an equilibrium strategy for the sellers to reveal as little information about the state of the world as possible.

1.5 Discussion and Conclusions

In this paper we have pointed out that there is a much broader scope for the common value model than conventional economic wisdom suggests. More precisely, a common value component can be expected to be present in any auction which fulfills the following two conditions: First, the good for sale must have some substitutes, i.e. buyers must have a relevant alternative (other than not buying at all) to winning the current auction. Second, there must be some uncertainty among the buyers about how valuable that alternative is.

Both conditions seem to be relatively weak, so the question is not really whether a common value component exists, but how strong it is. One area in which it should matter a lot are internet auction on platforms like ebay: Most of the objects traded there are everyday items, many of them even have perfect substitutes, because they are brand-new and have been mass produced. In addition, many buyers on ebay know

[17]There is a second important, albeit more subtle, difference between the two models: In the derivation of the Linkage Principle, it is assumed that the auctioneer reduces uncertainty about the good's value by making the states of the world more similar to each other. In terms of our model, she reduces the difference between d_h and d_l. By contrast, in our model the auctioneer can only provide the buyers with an additional signal about the state of the world, d_h and d_l remain the same as before.

[18]Mezzetti et al. (2004) find that a result similar to ours holds in a two-period model.

relatively little about the market in which they are active. Hence, a relatively strong common value component can be expected there even though the items for sale are mostly standard, and resale plays only a minor role. Since common value auctions differ in many respects from their private value counterparts (and it is relatively difficult to distinguish between them on empirical grounds[19]), this insight should be of interest to auction designers and bidders alike.

Of course, our model can be extended in a variety of ways. Its most apparent weakness lies in the very passive role of the sellers, and attempts to remove this shortcoming suggest themselves. However, we believe that there are two good reasons *not* to allow any seller to set a (secret) reserve price. First, it seems unlikely that doing so would yield any (substantial) new insights. Second, the incorporation of reserve prices would inevitably give rise to the following problem: If the sellers have enough commitment power to set a reserve price ex-ante, we run into a monopoly paradox a la Diamond (1971), i.e. in every equilibrium the sellers will - in both states - set a reserve price of v.[20] If, on the other hand, the sellers cannot commit to a reserve price (but only refuse to sell if the price is too low), the buyers' bidding strategy changes, because the threshold they have to pass in order to win the auction is no longer identical to the price they have to pay in case of a win. Thus, introducing reserve prices appears to be a major technical challenge. An alternative way of "animating" the sellers would be to let them decide whether or not they want to enter the market (at some cost). This would yield an additional equilibrium condition (the entrance cost of the marginal seller has to equal her expected revenue), but the model would not change otherwise.

Our result that it can never be optimal for a seller to provide losing bidders with any information about what happened in past auctions relies on the assumption that sellers cannot coordinate on a common informational regime (in which case they would

[19]The problem of how to distinguish between the private and common value models is, among others, discussed by Athey and Haile (2002)

[20]Note that this problem only arises, because buyers' valuations are all identical. The Diamond paradox immediately breaks down once the homogeneity assumption is dropped, see Lauermann (2006).

not take the distribution functions of buyers' types as given). One might therefore wish to relax that restriction.

In our opinion, however, the most promising avenue for future research is to use our model as a framework for more applied analyses of decentralized markets with aggregate uncertainty. As pointed out in the introduction we believe such uncertainty to play an important role in many markets, including those for housing and labor. Existing models which assume the distribution of types to be common knowledge may miss important aspects of economic behavior. For example, a considerable share of observed search activities in all kinds of markets can be attributed to agents trying to find out *what a good price is* rather than looking for one. Consequently, in order to improve our understanding of phenomena like frictional unemployment we need to get a better grasp of how markets are affected by aggregate uncertainty. It is our hope that the insights offered in the present paper will help us to achieve this goal.

Chapter 2

Status, Redistribution, and the Value of Consumption Signals

2.1 Introduction

Over the past couple of decades economists have accumulated compelling evidence that people do not only care about their own consumption, but also about the consumption of others. More specifically, while cross section studies consistently find that rich individuals tend to be happier than poor ones, even long periods of economic prosperity seem to leave a society's average happiness unaffected.[1] Hence, it appears that people compare themselves to their fellow citizens. Being richer than "the Joneses" makes them happy and vice versa. Technically speaking, most people seem to care about their respective rank in the income distribution, which one might call their personal *status*.

While this fact has been recognized for quite a while, it has only recently started to play a role in the analysis of redistributive politics. At first glance, one may ask what role it *can* play, because redistribution is usually done in an order-preserving way, so that nobody's rank in the income distribution changes. However, there are at least two ways of arguing that status considerations *do* have an impact on the consequences

[1] A comprehensive survey of this issue is provided by Frey and Stutzer (2002).

of redistribution.

One approach, due to Corneo and Grüner (2000), is based on the idea that the value of status stems from its signaling power. Consider, for example, a marriage market in which more and less attractive people look for a partner. A person's "matching value" cannot be observed, but is assumed to be correlated with her (gross) income, the rationale behind this assumption being that a number of skills (intelligence, eloquence, empathy etc.) are valued both on the labor market and on the marriage market.[2] While people's incomes are private information as well, this is not true of their consumption levels. Thus, marriages are arranged on the basis of consumption signals. Corneo and Grüner argue that the amount of observed consumption is likely to depend on some factors other than a person's income. Redistribution increases the relative importance of these "noise" factors, thereby changing the allocation of status.

A different line of argument is pursued by Hopkins and Kornienko (2004) who take up Hirsch's (1976) view that a person's status does not depend on her total consumption, but rather on how much of a certain "positional good" she consumes ("conspicuous consumption"). They consider a status game in which consumption decisions are taken strategically and show that redistribution leads to an increase in conspicuous consumption. The intuition behind this result is as follows: When income differences are large, the rich have to spend only a small fraction of their income on the positional good in order to set themselves apart from the poor. As income differences get smaller, social competition gets more intense, because the poor find it easier to overtake the rich. Since the total amount of status available is fixed, any increase in conspicuous consumption has a detrimental effect on welfare. Thus, Hopkins and Kornienko conclude that their model is able to explain the empirical observation that greater equality does not necessarily lead to a happier society.[3]

[2] It should be noted that, while wealth and attractiveness are positively correlated, there is no *causal* relationship between the two. Thus, taking away all the money from an initially rich person would not affect her attractiveness.

[3] There is a number of studies on this issue, and their results are very mixed. See, for example, Alesina et al. (2003) and Clark et al. (2005).

The persuasiveness of this argument partly depends on whether one views the value of status as being "hardwired" or "instrumental". If people aspire to higher status as an end in itself, it may be reasonable to assume that even very small differences in spending on the positional good can lead to vastly different levels of personal satisfaction (just like, in sports, very small differences in effort or talent can decide on victory and defeat).[4] Many economists, however, are reluctant to *directly* include social concerns as an argument in the utility function (Postlewaite, 1998). Instead, they argue that people care about their status, because a high social standing opens up additional consumption opportunities. More specifically, status may - as in the model of Corneo and Grüner - serve as a signal (of attractiveness, for example) according to which important non-market goods (like, for example, marriage partners) are allocated. Taking this line of reasoning seriously, the analysis of Hopkins and Kornienko only remains valid if one assumes a stable one-to-one relationship between income and conspicuous consumption. Since such an assumption essentially precludes any relevant heterogeneity of the agents (except for that in income), we find it rather hard to swallow.

In addition, casual evidence does not necessarily seem to suggest that people from Scandinavia (where income inequality is relatively small) spend a much larger share of their income on Mercedes cars and Rolex watches than those in Singapore or the United States (where income inequality is high compared to other developed countries). If anything, the converse appears to be true.

The purpose of this chapter therefore is to investigate some properties of a hybrid model which features both the concept of conspicuous consumption and the twin-assumption of (a) (exclusively) instrumental concerns for status and (b) a noisy relationship between income and status. The value of status in such a model is endogenous and we show it to depend on the income distribution. When income differences are large, the correlation between conspicuous consumption and income is almost perfect,

[4]The main argument in favor of direct rank concerns is that the survival of a person (in times of food shortage etc.) may crucially depend on her *relative* fitness. Thus, individuals who care about their status are likely to have a competitive advantage in the evolutionary process.

so high status guarantees an attractive partner. In contrast, when income differences are very small, the marriage market operates almost at random, so people care little about their status. Put differently, when (almost) everyone can afford a certain status good, nobody will buy it, because doing so would convey (almost) no information about the buyer's attractiveness.

Thus, our analysis adds a second, countervailing effect to the one identified by Hopkins and Kornienko: Redistribution does not only make it easier for the poor to compete with the rich (in the fight for status), it also reduces the value of status altogether. The relevant question is how strong these two effects are relative to each other. We show the relationship between income inequality and conspicuous consumption to be hump-shaped. When income differences are large, the Hopkins-Kornienko effect dominates, whereas the converse holds true in an egalitarian society. Hence, social competition is most intense at intermediate levels of income inequality.

At the heart of our analysis lies the assumption that a person's consumption pattern does not perfectly reveal her income. One can imagine a number of reasons for why this should be true, here we will mention just three: First, people generally differ with respect to their preferences. Some find it extremely important to be matched with an attractive partner, others do not care about this issue at all. Clearly, the former ones will tend to spend a higher fraction of their income on the positional good than the latter, thereby making themselves look more attractive than they actually are. Second, consumption signals may be noisy in the usual sense, because only some random fraction of an individual's (conspicuous) consumption is actually observed. Third, even when conspicuous consumption is fully observable, different people generally have to spend different amounts of money in order to acquire the same status. This fact is not only due to actual price dispersion,[5] but also stems from differences in people's ability to make others believe that they are rich. For example, some people are good at finding fashionable clothes even in rather cheap boutiques. Their ability to

[5]The importance of price dispersion has been recognized by economists at least since the publication of Varian's (1980) seminal paper.

identify goods with a favorable cost-performance ratio effectively makes these people face lower prices. Essentially, all three types of heterogeneity work into the same direction, and it is only for technical convenience that, henceforth, we are going to work with the third one.[6]

Our contribution differs from that by Hopkins and Kornienko not just in economic substance, we also employ different analytical tools. Whereas they model social competition as a first-price auction (and then exploit this formal analogy by using insights from auction theory), we resort to signaling techniques.

The remainder of this paper is organized as follows: In the next section we present a formal model of noisy signaling in a matching market. The effect of redistribution on the amount of conspicuous consumption is studied in Section 2.3. Section 2.4 summarizes and concludes.

2.2 The Model

2.2.1 Structure of the Model

We consider an endowment economy which is populated by a unit-mass continuum of agents. The agents engage in a two-stage game: At the first stage, they split their wealth between a normal and a positional good. Based on their consumption decisions and endowments, the agents then - at the second stage - get matched in pairs. The payoff of an agent depends both on how much of the two goods she consumes and on whom she gets matched with.

[6] In contrast to us, Corneo and Grüner employ the more classical second approach, and one may wonder why we do not follow them. The answer to this question is simple: Our model is more complicated than theirs in that we have to deal with strategic consumption choices. Thus, additional feasibility constraints arise which force us to work with the assumption of heterogeneous prices.

2.2.2 Agent i and her Problem

An agent $i \in [0,1]$ is characterized by a three-dimensional "character" vector $\omega_i = (v_i, w_i, p_i)$. We discuss the elements of this vector in turn: v_i is a measure of agent i's attractiveness, her so-called matching value. w_i tells us how much money (income or wealth) she is endowed with. The money can be divided between a normal good, c_i^n, and a positional good, c_i^p. The positional good does not directly yield any utility, but it relates to x_i, the matching value of agent i's partner. The third and final dimension along which the agents differ is the price they have to pay in order to obtain one unit of the positional good, p. The assumption of heterogeneous prices is certainly unorthodox, it can, however, be defended along the lines described in the introduction.[7]

All three elements of the character vector are continuously distributed random variables with support $[0, \infty)$. p is statistically independent of both v and w. In contrast, there is a strictly monotone and entirely deterministic relationship between v and w.[8] For the sake of tractability, we assume that agent i's preferences can be described by the log-linear utility function

$$V_i = v_i(x_i, c_i^n) = \ln x_i + \ln c_i^n \tag{2.1}$$

Normalizing the price of the standard good to one, her budget constraint is given by

$$w_i = p_i \cdot c_i^p + c_i^n \tag{2.2}$$

[7] An equivalent, more conventional (but notationally more burdensome) way of modeling heterogeneity with respect to the cost of status acquisition is to assume that the agents only buy a "raw material" which serves as an input in the production of the positional good. Assuming that the production technology is linear and that the agents differ with respect to their technology parameter, we get exactly the same results as with heterogeneous prices.

[8] Given the deterministic relationship between w and v, one may wonder why we consider these two variables separately. The reason is that we want redistribution (the effects of which are analyzed in Section 3) to affect the distribution of w, but *not* that of v.

Solving equation (2.2) for c_i^n, taking logs on both sides, and using the result to substitute for $\ln c_i^n$ in (2.1), we obtain the following description of agents' preferences over bundles of c^p and x.

$$U_i = \ln\left(\frac{w_i}{p_i} - c_i^p\right) + \ln x_i + \ln p_i \tag{2.3}$$

Equation (2.3) reveals that, while an agent's utility is a function of both w and p, these two parameters solely interact with c^p through the ratio w/p. Therefore, any agent's choice of c^p should only depend on $\theta \equiv w/p$. In the jargon of information economics, w and p can be *aggregated* into θ. The intuition behind this insight (which we will take up in the next subsection) consists of two steps: First, two agents with identical θ have to give up the same *proportion* of their normal consumption (c^n) in order to gain a one-point increase in conspicuous consumption (c^p). Second, as we work with a log-linear utility function, any decrease in c^n by the same *proportion* translates into a decrease in U by the same absolute *amount*. Thus, the two agents face exactly the same incentives.

The observation that an agent's behavior is essentially driven by a one-dimensional variable allows us to cast the model into a relatively standard signaling framework. Let $r : \mathbb{R}_+ \to \mathbb{R}_+$ denote the *response function*, i.e. a function which tells us how attractive a partner the agents expect to get in response to a given amount of conspicuous consumption, $\bar{x} \equiv \exp\left(E\left[\ln x\right]\right) = r(c^p)$.[9] While the response function is endogenous to our model, all agents take it as given. Thus, they maximize the function

$$U = u\left(c^p, \theta, r(c^p)\right) = \ln\left(\theta - c^p\right) + \ln r(c^p) \tag{2.4}$$

where the (constant) term $\ln p_i$ has been dropped for convenience. In the next subsection we will use function (2.4) to characterize the model's equilibrium.

[9] In the present context the term "expected" refers to the geometric (rather than the arithmetic) mean, because the agents' utility function is log-linear (rather than simply linear).

2.2.3 The Equilibrium of the Model

We begin our analysis of the model at the matching stage. The marriage market we consider operates under the assumption that the formation of a partnership requires both partners to consent. Thus, matchings must be stable in the sense of Gale and Shapley (1962).[10] Under complete information, this would require all couples to consist of equally attractive agents. (Otherwise, the more attractive individual would have an incentive to break up.) In our game, however, the agents' characteristics (as well as their spending on the normal good) are assumed to be unobservable, so we require partners to be identical in terms of $r(c^p)$, i.e. they must be equally "promising". In addition, if the function $r(\cdot)$ takes on the same value at different levels of c^p, only the lowest of these levels will be chosen by any agent. Thus, in any equilibrium all partners must have consumed exactly the same amount of the positional good.

Let us now turn to the first stage: As emphasized above, the marginal utility of an increase in c^p is determined by θ only, neither does it depend on v nor on the decomposition of θ into its components w and p. Therefore, we refer to θ as an agent's *type* and posit that her strategy, denoted as $s : \mathbb{R}_+ \to \mathbb{R}_+$, is a function of θ only, $c^p = s(\theta)$. This reduction in dimensionality simplifies our analysis of the model immensely. Nevertheless, we are faced with a problem to which most signaling games are prone: multiple equilibria can arise. In what follows we focus on fully separating equilibria, i.e. on equilibria in which the agents use strictly increasing strategies,[11] so that an agent's action fully reveals her type. This assumption gives rise to the following equilibrium definition:

Definition 2 *(Weak perfect Bayesian equilibrium in strictly increasing strategies): An equilibrium of the game consists of a response function $r(c^p)$ and a strategy $s(\theta)$ such that*

[10] A more formal description of a matching market very similar to ours can be found in Corneo and Grüner (2000) and Cole et al. (1992).

[11] In addition we impose the standard assumption that the agents' strategies are pure and symmetric.

1. *The agents have rational expectations about the attractiveness of their partner, i.e.* $r(s(\theta)) = \exp(E[\ln(v|\theta)])\ \forall\theta$

2. *The strategy is utility maximizing for all agents, i.e. for all* θ *we have that* $U(s(\theta), \theta, r(s(\theta))) \geq U(c^p, \theta, r(c^p))\ \forall c^p$

While Definition 2 tells us what an equilibrium of our game (in strictly increasing strategies) should look like, we still need to find out whether such an equilibrium exists at all. To this end, let us first have a closer look at how θ and $\bar{x}$ relate to each other. As can be seen from Definition 2, this relationship - denoted as $\bar{x} = h(\theta)$ in what follows - is exogenously given by $E[\ln(v|\theta)]$, i.e. we do not need to know the model's equilibrium in order to determine the expected attractiveness of an agent's partner. This allows us to directly make assumptions on $h(\cdot)$. What we need to ensure is that high-type agents get some return on their investment in the positional good (for they would otherwise choose to save that money). Technically speaking, we need to impose the following

Assumption 1: $E[\ln v|\theta_1] > E[\ln v|\theta_2]$ for all $\theta_1 > \theta_2$, i.e. high types are - on average - always more attractive than low ones.

In addition, we need to make sure that $(dc^p/d\bar{x})_{\bar{u}}$, the marginal rate of substitution between an agent's spending on the positional good and the attractiveness of her partner, is increasing in θ. This condition - which is nothing but the well-known single-crossing condition (s.c.c.) - needs to hold, because otherwise low-type agents would not necessarily suffer more from (increases in) conspicuous consumption than do high-type agents. Thus, both types would have an incentive to mimic each other. In our model the s.c.c. holds, because the derivative of U with respect to c^p is - as equation (2.3) reveals - increasing in θ.

The s.c.c. does not only imply that a strictly separating equilibrium exists. Mailath (1987) has also shown that, in signaling games with a continuous type space (like the one considered here), the set of strictly separating equilibria contains just one element: This is the so-called Riley-equilibrium (Riley, 1979) in which all agents spend just enough on the positional good to make it unprofitable for lower types to imitate them. Put differently, all agents are indifferent towards a marginal increase or decrease in spending on the positional good. Hence, we can characterize the model's equilibrium by going back to equation (2.4), taking the derivative of the agents' utility function with respect to c^p, setting it to zero, and then solving for $r'(c^p)$. This gives us the following implicit expression of an agent's indifference point:

$$r'(c^p) = r(c^p)/\left(\theta - c^p\right) \tag{2.5}$$

where $r(c^p) = 0$. The solution to this (linear first order) differential equation can easily be found by applying the appropriate formula.[12] It is given by

$$r(c^p) = \exp\left(\int_0^{c^p} \frac{1}{\theta - t} dt\right) \tag{2.6}$$

Since $r(\cdot)$ and $s(\cdot)$ are interlinked by the identity $s(\theta) = r^{-1}(h(\theta))$, equation (2.6) fully characterizes the equilibrium of the model.

2.2.4 What Drives Social Competition?

The intensity of social competition can best be measured by the level a given type's spending on the positional good. This is why, in what follows, our interest focuses on the agents' strategy, i.e. on the question of how c^p depends on θ. In order to answer this question we need to carry out some manipulations of equation (2.5). Recall that $r(c^p) = h(s^{-1}(c^p))$. Taking the derivative of this function with respect to c^p we get that $r'(c^p) = h'(\theta)/s'(\theta)$. Substituting for $r'(c^p)$ and $r(c^p)$ in (2.5) and solving for $s'(\theta)$ yields

$$s'(\theta) = \frac{(\theta - s(\theta)) \cdot h'(\theta)}{h(\theta)} \tag{2.7}$$

In principle, this differential equation can be solved in the same way as (2.5). Rather than doing so, however, we will analyze it directly. To this end, consider the following reformulation:

[12] The formula can, for example, be found in Simon and Blume (1994, p. 639).

$$c^p = s(\theta) = \int_0^\theta \frac{(t - s(t)) \cdot h'(t)}{h(t)} dt = \int_0^\theta (1 - \lambda(t)) \cdot \varepsilon(h(t), t) dt, \tag{2.8}$$

where the elasticity $\varepsilon(\bar{x}, \theta) \equiv \frac{d\bar{x}}{d\theta} \cdot \frac{\theta}{\bar{x}}$ indicates how strongly θ reacts (in equilibrium) to a change in $\bar{x}$. A large value of $\varepsilon(\bar{x}, \theta)$ means that small percentage changes in θ are sufficient to induces large ones in $\bar{x}$, i.e. relatively similar types of agents enjoy very different expected matching values. $\lambda(\theta)$ is defined as $\lambda(\theta) \equiv \frac{s(\theta)}{\theta}$. We are now ready to state our first lemma:

Lemma 2. Consider two economies, indexed by 1 and 2. If $\varepsilon_1(\bar{x}, \theta) < \varepsilon_2(\bar{x}, \theta) \; \forall \; \theta$, then $s_1(\theta) \leq s_2(\theta) \; \forall \; \theta$.

Proof. Note first that $s(\cdot)$ is a continuously differentiable function and we have $s_1(0) = s_2(0) = 0$. Thus, if $s_1(\theta)$ is ever to exceed $s_1(\theta)$, there must be a region in which (a) $s_1(\theta) > s_2(\theta)$ and (b) $s_1'(\theta) > s_2'(\theta)$. However, condition (a) implies that $1 - \lambda_1(\theta) < 1 - \lambda_2(\theta)$. Since we also have that $\varepsilon_1(\bar{x}, \theta) < \varepsilon_2(\bar{x}, \theta)$, $s_1'(\theta)$ must - by equation (2.8) - be *smaller* than $s_2'(\theta)$. Thus, conditions (a) and (b) cannot be fulfilled simultaneously. ■

The intuition behind Lemma 2 is as follows: When similar types of agents get partners with very different (expected) matching values (i.e. $\varepsilon(\bar{x}, \theta)$ is large), low types find it very attractive to disguise themselves as high ones. In order to avoid being imitated (and thus being deprived of their attractive partners), the high types need to spend a large part of their wealth on the positional good. Conversely, when it is either too unimportant or too difficult for the poor to outdo the rich, conspicuous consumption is low.

While Lemma 2 makes clear that the intensity of social competition depends on $\varepsilon(\bar{x}, \theta)$, we do not know yet how this elasticity is affected by redistribution. This question will be addressed in Section 2.3.

2.3 The Effects of Redistribution

2.3.1 Distributional Assumptions

The proof of Lemma 2 only relies on v, w and p being continuously distributed, it does not require any specific distributional assumptions. By contrast, in the present section we will assume that the character vector follows a multivariate log-normal distribution with parameters σ_w^2, σ_v^2, σ_p^2, μ_w, μ_v and μ_p. The latter two are normalized to zero, $\mu_v = \mu_p = 0$, whereas μ_w may be positive or negative, $\ln w \sim N(\mu_w, \sigma_w^2)$. As mentioned before, there is perfect correlation between $\ln w$ and $\ln v$, whereas p is independent of the two, $\rho_{wp} = 0$ and $\rho_{vp} = 0$.[13]

2.3.2 Finding a Parametric Expression for $\varepsilon(\bar{x}, \theta)$

Our main goal in this subsection is to express $\varepsilon(\bar{x}, \theta)$ in terms of σ_w^2, σ_v^2, σ_p^2 and μ_w. To this end, we first need to know how θ and $\bar{x}$ are distributed. This is the subject of our next lemma.

Lemma 3: $\ln \theta$ is normally distributed with mean μ_w and variance $\sigma_w^2 + \sigma_p^2$, while $\ln \bar{x}$ is normally distributed with mean zero and variance $\sigma_w^2 \cdot \sigma_v^2 / (\sigma_w^2 + \sigma_p^2)$.

Proof. The proof of the Lemma's first part is straightforward: It directly follows from the fact that both w and p follow a log-normal distribution and $\theta = \frac{w}{p}$, As for $\ln \bar{x}$, note first that, since there is a one-to-one relationship between v and w, $\ln v = \frac{\sigma_v}{\sigma_w} \cdot (\ln w - \mu_w)$. Thus, we have that $\ln \bar{x} = \frac{\sigma_v}{\sigma_w} E[\ln w | \theta]$. As w and p are both log-normally distributed, we know what the expectation of $\ln w$ conditional on θ looks like (see, for example, Mood, Grabill, and Boes, 1974, pp.167-168).

$$E[\ln w | \theta] = \mu_w + \frac{\sigma_w^2}{\sigma_w^2 + \sigma_p^2} \cdot (\ln \theta - \mu_w) \tag{2.9}$$

Recall next that, when z is a random variable and both a and b are scalars, we have that

$Var[a \cdot (z - b)] = a^2 \cdot Var[z]$ and $E[a \cdot (z - b)] = a \cdot (E[z] - b)$. Thus, from

$$E[\ln x | \theta] = \frac{\sigma_v \cdot \sigma_w}{\sigma_w^2 + \sigma_p^2} \cdot (\ln \theta - \mu_w) \tag{2.10}$$

---

[13] Note that these distribution functions satisfy Assumption 1.

we get that $\ln \bar{x} \sim N\left(0, \sigma_v^2 \cdot \frac{\sigma_w^2}{\sigma_w^2+\sigma_p^2}\right)$, because $\ln \theta \sim N(\mu_w, \sigma_w^2 + \sigma_p^2)$. ∎

Note that the distribution of $\ln \bar{x}$ does not depend on μ_w in any way. The expression for the variance of $\ln \bar{x}$ can intuitively be understood as follows: when an agent observes a high value of θ, she presumes that both w and p account for part of it. How large a share she attributes to w depends on the relative variances of w and p. When σ_p^2 is large relative to σ_w^2, the variance of $\ln \bar{x}$ is small, because the signals provided by θ (regarding v) are so noisy that it would be irrational for an agent to attach strong meaning to them. The best she can do is suppose that $\ln \bar{x}$ lies somewhere close to the expectation of $\ln v$. Conversely, when σ_p^2 is much smaller than σ_w^2, the correlation between $\ln \theta$ and $\ln v$ is almost perfect, so the variance of $\bar{x}$ is only slightly smaller than that of v.

The next step is to find out what the function $h(\cdot)$ looks like, i.e. how θ relates to $\bar{x}$. Since there is a strictly monotone and positive relationship between the two, we can construct $h(\cdot)$ as follows: Let the cumulative density functions (c.d.f.) of θ and $\bar{x}$ be denoted by $F_\theta(\cdot)$ and $F_{\bar{x}}(\cdot)$, respectively. The functional relationship between θ and $\bar{x}$ is then implicitly defined by $F_\theta(\theta) = F_{\bar{x}}(\bar{x})$, i.e. $\theta = F_\theta^{-1}(F_{\bar{x}}(\bar{x}))$. Letting Φ denote the c.d.f. of the standard normal distribution and recalling that both θ and $\bar{x}$ are log-normally distributed, we thus have that

$$\bar{x} = h(\theta) = \exp\left(\sigma_{\bar{x}} \cdot \Phi^{-1}\left(\Phi\left(\frac{\ln \theta - \mu_w}{\sigma_\theta}\right)\right)\right) = \exp(\mu_w)^{-\frac{\sigma_{\bar{x}}}{\sigma_\theta}} \cdot \theta^{\frac{\sigma_{\bar{x}}}{\sigma_\theta}} \tag{2.11}$$

The elasticity $\varepsilon(\bar{x}, \theta)$ is found by taking the derivative of this expression (with respect to θ), multiplying it by $\frac{\theta}{\bar{x}}$ and substituting for $\bar{x}$ using (2.11). This yields:

$$\varepsilon(\bar{x}, \theta) = \frac{\sigma_{\bar{x}}}{\sigma_\theta} \cdot \theta^{\frac{\sigma_{\bar{x}} - \sigma_\theta}{\sigma_\theta}} \cdot \frac{\theta}{\bar{x}} = \frac{\sigma_{\bar{x}}}{\sigma_\theta} \tag{2.12}$$

This result is most easily understood by bringing to one's mind the fact that $\varepsilon(\bar{x}, \theta) = \frac{\partial \ln \bar{x}}{\partial \ln \theta}$. Both $\ln \theta$ and $\ln \bar{x}$ are normally distributed (and there is a one-to-one relationship between the two), so it follows that $\Phi(\ln \bar{x}) = \Phi\left(\sigma_{\bar{x}}/\sigma_\theta \cdot (\ln \theta - \mu_w)\right)$. Thus, in response to every change of $\ln \theta$ by one unit, $\ln \bar{x}$ must change by $\sigma_{\bar{x}}/\sigma_\theta$ units.

2.3.3 Redistribution and Conspicuous Consumption

We model redistribution as a change in σ_w. Since the distribution of income is assumed to be *log*-normal (and not normal), this entails two problems. First, an isolated change in σ_w does not leave the expectation of w unaffected. Thus, every increase in σ_w needs to be compensated by a decrease in μ_w (because we want redistribution to be budget-neutral). Normalizing average income in the economy to unity, we have that $\mu_w = -\frac{\sigma_w^2}{2}$.[14] However, since $\varepsilon(\bar{x}, \theta)$ is independent of μ_w, this does not cause any real difficulties. Second, the redistribution-scheme we consider is non-linear, i.e. it cannot be interpreted as a proportional income tax whose revenues are spread equally over all citizens. However, neither do we have the impression that linear schemes prevail in the real world nor does it seems to us that the assumption of non-linear redistribution (which is for mathematical convenience only) drives any of the paper's qualitative results.

As it is apparent from Section 2.3.2, both σ_θ and σ_s depend on σ_w. Moreover, in both cases the effect is positive. A high degree of income inequality implies a large variance of the type distribution, but it also leads to very pronounced social stratification (i.e. a very close relationship between w and $\bar{x}$). The following central proposition of this paper sheds light on the question of how strong these two effects are relative to each other.

Proposition 9: $\partial s(\theta, \sigma_w)/\partial \sigma_w < 0 \ \forall \ \theta$ if $\sigma_w > \sigma_p$ and $\partial s(\theta, \sigma_w)/\partial \sigma_w > 0 \ \forall \ \theta$ if $\sigma_w < \sigma_p$, i.e. a reduction in income inequality may - depending on the relative size of σ_w and σ_p - either increases or decrease a given type's spending on the positional good.

Proof. From Sections 2.2.4 and 2.3.2 we know that

$$\varepsilon(\bar{x}, \theta) = \frac{\sigma_{\bar{x}}}{\sigma_\theta} = \frac{\sigma_v \cdot \sigma_w}{\sigma_w^2 + \sigma_p^2} \tag{2.13}$$

[14] Recall that the expectation of a log-normally distributed random variable z with parameters μ and σ^2 is given by $E[z] = \exp\left(\mu + \frac{\sigma^2}{2}\right)$

Simply taking the derivative of this expression with respect to σ_w yields:

$$\frac{\partial \varepsilon(\bar{x}, \theta)}{\partial \sigma_w} = \frac{\sigma_v \cdot \sigma_w^2}{\sigma_w^2 - \sigma_p^2} \tag{2.14}$$

Since σ_v and σ_w are both positive, this expression is positive if and only if σ_w is bigger than σ_p. Thus, equation (2.14) proves our claim. ■

Proposition 9 shows that redistribution (from the rich to the poor, i.e. a *decrease* in σ_w) leads to an increase in the intensity of social competition when the variance of the wealth distribution is larger than that of the price distribution. Yet, it *mitigates* social competition when the converse holds true. Thus, the relationship between economic inequality and social competition is not monotone, but hump-shaped. This result stands in contrast to the findings of Hopkins and Kornienko who argue that redistribution *always* intensifies social competition. Yet, we are able to obtain an even stronger result:

Proposition 10: For all $\sigma_p > 0$, all agents' spending on the positional good goes to zero as differences in income vanish, i.e. $\lim_{\sigma_w \to 0} s(\theta) = 0 \; \forall \; \theta$.

Proof. Key to the proof is the observation that we know from Lemma 3 that $\lim_{\sigma_w \to 0} \sigma_{\bar{x}} = 0$, whereas $\lim_{\sigma_w \to 0} \sigma_\theta > 0$. As it is apparent from equation (2.11), this implies that $h'(\theta) = 0 \; \forall \; 0 < \theta < \infty$, i.e. almost all agents get equally attractive partners. Hence, the integrand in equation (2.7) is equal to zero except at $\theta = 0$ which proves the proposition. ■

The message contained in Proposition 10 is a strong one: When all agents are endowed with the same amount of money, social competition completely disappears. The intuition behind this result is clear: It is the correlation between θ and v that makes the agents consume the positional good. As this correlation goes down to zero (because the variation in w is entirely dominated by that in v), consumption signals do not contain any relevant information anymore. Thus, the agents stop sending them.

2.3.4 Redistribution and Welfare

Our main goal in the present subsection is to analyze the effect of redistribution on U_i, the individual welfare of our exemplary agent i. Generally, redistribution affects U_i

through two different channels: It changes the distribution of types in the population (thereby affecting $g(\cdot)$), but also i's own income w_i. In principle, these two effects could be analyzed separately. However, we will link them together by taking i's rank in the income distribution as fixed.[15] Put formally, we assume that $\ln w_i - \mu_w$ is proportional to σ_w. This allows us to make the following statement about the effect of redistribution on agent i's partner:

Proposition 11: Redistribution (from the rich to the poor) enhances the attractiveness of agent i's partner iff $\ln p_i \cdot (\sigma_p^2 - \sigma_w^2) - 2 \cdot (\ln w_i - \mu_w) \cdot \sigma_p^2 > 0$

Proof. Recall from Lemma 3 that the expected log-attractiveness of agent i's partner is given by $\ln \bar{x} = \sigma_v \cdot \sigma_w / (\sigma_w^2 + \sigma_p^2) \cdot (\ln \theta - \mu_w)$. We expand this formula by a proportionality constant, denoted as k, in order to be able to study simultaneous changes in σ_w and $\ln w_i - \mu_w$. This gives us

$$\ln s(k) = \frac{\sigma_v \cdot \sigma_w \cdot k \cdot (k \cdot (\ln w - \mu_w) - \ln p)}{k^2 \cdot \sigma_w^2 + \sigma_p^2} \tag{2.15}$$

Taking the derivative of (2.15) with respect to k and then setting k to one (which is without any loss of generality) yields the condition named in the Proposition. ■

Proposition 11 is easiest to interpret when $\sigma_w^2 = \sigma_p^2$. In that case redistribution increases $\ln \bar{x}_i$ (the expected matching value of i's partner) if $\ln w_i$ lies below average, i.e. if i belongs to the poorer half of the population. While the intuition that redistribution should be beneficial to low-income agents remains correct when $\sigma_w^2 \neq \sigma_p^2$, the total effect of a change in σ_w then also depends on p_i. When $\sigma_p^2 > \sigma_w^2$, redistribution tends to make agents who face a high p better off, because it reduces the variance of $\bar{x}$. When $\sigma_p^2 < \sigma_w^2$, on the other hand, agents who enjoy favorable prices gain because their rank in the distribution of θ improves. When $\sigma_w^2 = \sigma_p^2$, these two effects just balance off each other.

Let us now turn to the effect of redistribution on c_i^n. Here, the following lemma will prove helpful:

[15] An alternative approach - pursued by Hopkins and Kornienko - is to fix i's actual income (rather than her rank in the income distribution).

Lemma 4: $s'(\theta) < 1$, i.e. there cannot be pairs of agents such that (after adjusting for differences in p) $c_1^p > c_2^p$ and $c_1^n < c_2^n$.

Proof. The proof is similar in style to that of Lemma 2: Since $s'(\cdot)$ is a continuous function, there must, in order for $s'(\theta)$ to get bigger than one, a region in which (a) $s'(\theta) > 1$, i.e. $\lambda(\theta) < 1 - \varepsilon$ and (b) $s''(\theta) = 1/\theta \cdot 1/\varepsilon \cdot (\lambda(\theta) - s'(\theta)) > 0$. Merging these two conditions, we get that $1/\theta \cdot (1/\varepsilon \cdot (1 - s'(\theta)) - 1 > 0$. As $s'(\theta) > 1$ (by assumption) this inequality cannot hold. Thus, we have a contradiction which means that conditions (a) and (b) can never be satisfied simultaneously. ■

What we learn from Lemma 4 is that increasing an agent's type by one unit leads to an increase of that agent's spending on the positional good by *less* than one unit. This insight plays an important role in the proof of the following proposition:

Proposition 12: When $\sigma_p^2 > \sigma_w^2$ all agents with $w < \exp(\mu + \sigma_w^2)$ gain from redistribution in terms of c^n. Conversely, all agents with $w > \exp(\mu + \sigma_w^2)$ suffer from redistribution when $\sigma_p^2 < \sigma_w^2$.

Proof. We first show that the population group named in the Proposition gains from redistribution in terms of *income.* To this end, consider an agent with income $\exp(\mu + k \cdot \sigma_w)$. By definition (see Section 2.3.3) this amount is equal to $\exp(-\sigma_w^2/2 + k \cdot \sigma_w)$. Thus, redistribution increases the agent's income if $k < \sigma_w$. What remains to be proven is that (if $\sigma_p^2 > \sigma_w^2$) every rise in income is mirrored by an increase in non-conspicuous consumption. Consider an agent whose income - in response to redistributive action - changes from w_1 to w_2, where $w_2 > w_1$. As p is constant the rise in w implies that $\theta_2 > \theta_1$. By Lemma 2, we have that $s_2(\theta) < s_1(\theta) \; \forall \; \theta$, i.e. spending of a *given* type is lower after redistribution than before. In addition (as we have just shown) $s'(\theta) < 1$, so it holds that $\theta_2 - s_2(\theta_2) > \theta_1 - s_2(\theta_1) > \theta_1 - s_1(\theta_1)$. Since $x = p \cdot (\theta - y(\theta))$ (where p is a constant unaffected by redistribution), we can conclude that $x_2 > x_1$. The second part of the proposition can be proven analogously. ■

Propositions 11 and 12 have an important political-economic interpretation. If redistribution alleviates the intensity of social competition, a sizeable share of the population (which may - depending on the parameter constellation - well constitute a majority) should be willing to support it.

So far we have focused on the utility of single agents, our next goal is to say something about the effect of redistribution on *social* welfare. As we want our model to be robust to monotone transformations of the utility function $u(\cdot, \cdot)$ (i.e. we do not want to rely on cardinal concepts of utility), it is impossible to set up a utilitarian welfare function. However, since the status game we consider is zero-sum, the total amount of spending on the positional good, denoted as $\Omega(\sigma_w)$ in the following, can be used as a measure of distance from Pareto optimality, which in turn may serve as an index of social welfare. Unfortunately, it is not admissible to directly interpret Proposition 9 as a statement about $\Omega(\cdot)$, because redistribution does not only affect $\varepsilon(\theta, \bar{x})$, it also changes the distribution of types in the population. (We will return to this issue in Section 2.3.5.) In contrast, Proposition 10 directly implies that total spending on the positional good goes to zero as the income distribution degenerates. Thus, in our model (where social competition is the only source of economic inefficiency), a Pareto optimum can be achieved by means of complete redistribution.

2.3.5 Numerical Evidence on the Effects of Redistribution

In Sections 2.3.3 and 2.3.4 we have offered several results on how redistribution affects individual well-being and the intensity of social competition. A number of important questions, however, had to be left open, because they are not susceptible to theoretical analysis. The purpose of the present subsection is to shed light on them by numerical means. We do not aim to calibrate the model, though, our goal simply is to solve it for different values of σ_w (ranging from 0.1 to 10), thereby providing evidence (albeit no definite proof) on the effects of redistribution. In order to be able to relate our results to those of the previous subsections, we will stick to the assumption of log-normality. For simplicity we set both σ_v and σ_p to one.

The first question we want to explore is how redistribution affects aggregate consumption of the positional good, i.e. what the function $\Omega(\sigma_w)$ looks like. Figure 3 reveals that (in qualitative terms) the relationship between these two variables differs hardly from that between σ_w and individual consumption of the positional good. Thus, the fact that redistribution changes the distribution of types does not appear to have any major effects on the shape of $\Omega(\sigma_w)$.

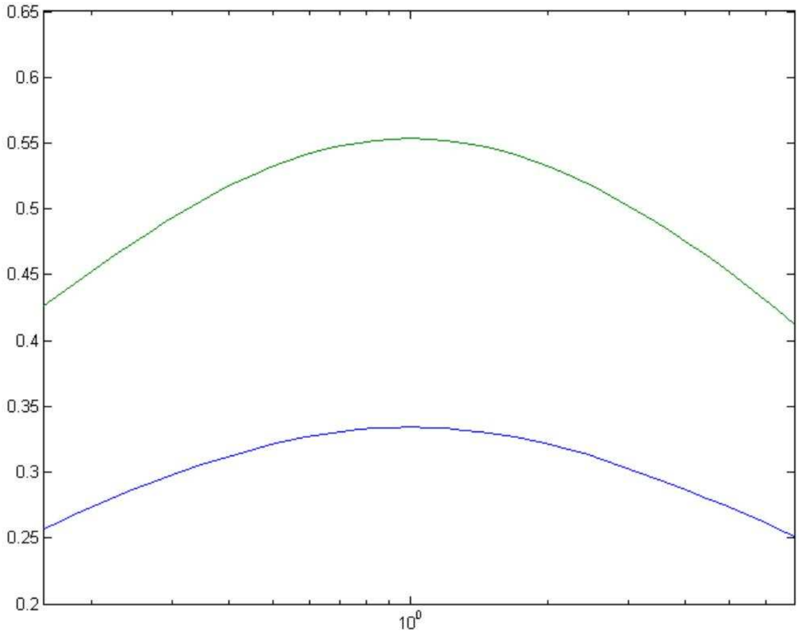

Figure 3: Relationship between σ_w and conspicuous consumption

Lower graph: Relationship between σ_w^2 (x-axis) and the conspicuous consumption of an agent with $p = w = 0$ (y-axis). Upper graph: Relationship between σ_w^2 and society's *aggregate* consumption of the positional good. Note that - in line with Proposition 3 - both graphs peak (approximately) at $\sigma_w = \sigma_p = 1$.

At the individual level, we would like to know which agents benefit / suffer from redistribution. Figure 4 depicts the effect of changes in σ_w on the utility level of agents with different levels of income.

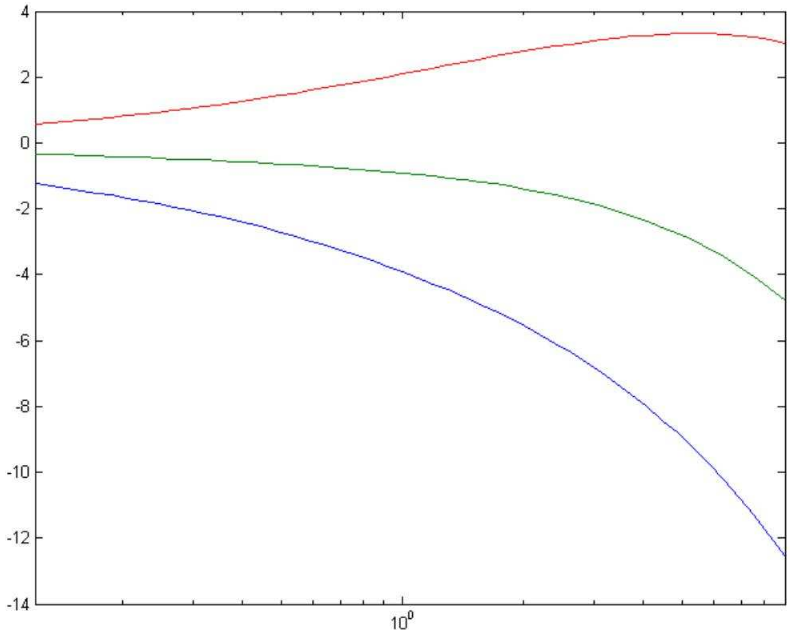

Figure 4: Relationship between σ_w and agents' utility

Lower graph: Relationship between σ_w^2 (x-axis) and the utility of an agent with $\ln w = \mu - 2 \cdot \sigma_w$ (y-axis). Middle graph: Relationship between σ_w^2 and the utility of an agent with $\ln w = \mu$. Upper graph: Relationship between σ_w^2 and the utility of an agent with $\ln w = \mu - 2 \cdot \sigma_w$. All three agents face average prices, $p = 1$.

Remarkably, even high-income agents suffer from increases in σ_w when σ_w already is relatively large. This is because (as mentioned in Section 2.3.3) we compensate for every increase in σ_w by a decrease in μ_w, and that decrease is proportional to σ_w^2 (whereas any agent's comparative income advantage is only proportional to σ_w). As σ_w approaches infinity, the mass of agents with income above any $\varepsilon > 0$ thus goes to zero. Therefore, most agents benefit from decreases in σ_w even when that decrease entails an increase in conspicuous consumption.

2.3.6 Testing the Model

Our model predicts that redistribution (from the rich to the poor) may well alleviate the intensity of social competition. This prediction is contradictory to the conclusions of Hopkins and Kornienko. In the present section we discuss how the implications of both models could be tested empirically.

In principle, it is easy to regress the amount of conspicuous consumption (per capita) on some measure of income inequality (like, for example, the Gini-coefficient). Doing so, however, requires one to somehow distinguish between "positional" and "non-positional" goods and to measure spending on them separately. At first glance, this problem may appear almost impossible to solve, but fortunately there is a nice way to get around them: One can simply use leisure (arguably the least conspicuous good) as a proxy for non-positional consumption and take all monetary income (as measured by the GNP) to be observable.[16] The number of annual working hours (divided by the total number of hours per year) can then be interpreted as the share of (potential) income spent on positional goods, i.e. it can be used as a measure of conspicuous consumption.

This kind of analysis could be based on either time-series or cross-section data. Both kinds of data have certain advantages: Time-series data do not suffer from (essentially unobservable) cultural differences between countries which may well confound the relationship between inequality and conspicuous consumption. They can, on the other hand, be subject to common time trends which may induce spurious correlations. Ideally, one should use panel data and control for both time trends and country-specific intercepts by means of dummy variables.

Unfortunately, we do not have enough data to conduct a formal analysis, but the available evidence is highly suggestive: Over the past 30 years the United States have experienced a marked increase in income inequality (Katz and Autor, 1999),

[16] Of course, not all income is consumed, some of it may also be invested (which is not possible in our model). However, investment (and net lending) can be regarded as measures of future consumption, so all we do is to interpret dynamic data in a static way.

while annual working hours have gone up steadily (Schor, 1992). By contrast, in most European countries income inequality has stayed roughly constant, and annual working hours have declined considerably. Thus, it appears that redistribution indeed alleviates the intensity of social competition. This (preliminary) finding corroborates our analysis while it contradicts the one of Hopkins and Kornienko.

2.4 Summary and Conclusions

We have shown that redistribution has two opposing effects on the intensity of social competition. On the one hand, the decrease in income differences makes it harder for the rich to set themselves apart from the poor. Thus, the former need to increase their investment into status goods in order to preserve their position as the society's "elite". However, as redistribution renders the relationship between a person's attractiveness and her status increasingly noisy, it also reduces the value of a high social standing. Thus, investment into status goods gets reduced. The latter effect dominates in egalitarian societies, whereas the converse hold true when income inequality is high. Hence, the relationship between the extent of redistribution and the intensity of social competition is hump-shaped.

The proofs of all propositions and lemmas in Section 2.3 take w, v, and p to be log-normally distributed. One important question is to what extent this very specific assumption drives our results. Fortunately, our propositions all have a very intuitive economic interpretation, so there is reason to believe that they are valid under much more general conditions than the ones of our model. Nevertheless, it seems prudent to reflect on "what can go wrong". In our view the most critical assumption is that $\ln p$ has unbounded support. In particular, when the support of $\ln p$ is bounded from *above* (i.e. an agent's type puts an upper bound on her attractiveness), there are always some low-type agents who never get matched to an attractive partner. All agents try to avoid being associated with this group of the "outcast". At the same time, its "members" can (in the presence of a linear redistribution scheme, for example) get relatively wealthy. In that case social competition remains at a high level even as income inequality converges to zero. This effect can occur regardless of how w is distributed. Conversely, when the support of $\ln p$ is bounded from *below*, there is a group of "elite" agents who always have an attractive partner. However, the size of this group shrinks to zero as income differences vanish, and it does not have an influence on the consumption behavior of the masses.

As emphasized in the previous subsection, redistributive politics affect a society's well-being through many channels, our model entirely focuses on just one of them. Its results should therefore be interpreted with care, we certainly do not plead for radicalness in the taxation of income. However, our finding that egalitarianism generally induces *low* levels of social competition may help to explain why equality ("égalité") is widely, though not universally, regarded as a value. In fact, one could go so far as to say that our model captures the essence of the socialist "Utopia":[17] The creation of a classless society in which people do not only get according to their needs, but also - rather than constantly struggling against each other - live together in peace and harmony.

[17]The imaginary island of "Utopia" was described by Thomas More in his classic treatise on the organization of human societies.

Chapter 3

Can we Trust Structural Estimates of the Returns to Schooling?

3.1 Introduction

During the past couple of decades very few economic parameters have attracted as much attention as the effect of attending school on subsequent earnings, the "returns to schooling". There are two reasons for the strong - and continued - interest in this issue. First, the returns to schooling appear to be an important variable for both individual decisions and - presuming that earnings are positively related to productivity - economic policy (Griliches, 1970). Second, it is so difficult to estimate the returns to schooling that we still do not know very well how large they actually are. The main econometric challenge lies in the endogeneity of the explanatory schooling variable: As people, rather than being randomly assigned a certain level of education, decide themselves when to leave school, those who go there for only, say, nine years may well be entirely different from those who complete college. Hence, it is not clear whether the observed positive relationship between schooling and earnings is *causal* or *spurious*, i.e. whether schooling has an effect on earnings or whether both variables are driven by a third, unseen factor.

In principle, there are two ways to overcome this problem. The first one is to look for a "random component" in individual schooling choices, i.e. a variable which influences the timing of school exit, but is unrelated to any skills which might be valued on the labor market. If such a variable (an "instrument") is available, one can consistently estimate the returns to schooling by instrumental variable (IV) techniques.[1] Finding reliable instruments, however, has turned our to be an extremely ambitious quest: Most candidate instruments are either "weak" (in the sense that they only have a very small bearing on schooling decisions)[2] or they are still related to agents' personal characteristics and, hence, must be counted on as having something to do with their earnings potential. In sum, while the IV approach has delivered a number of insights into the relationship between schooling and earnings, it is still met with (possibly justified) skepticism by many economists.

Building on work by Keane and Wolpin (1997), Belzil and Hansen (2002) have therefore proposed a life cycle model of sequential schooling decisions to estimate the returns to education. As we shall see, it is not easy to give a non-technical account of how their approach works. One of its central ingredients, however, is to relate the benefits of education to its costs. More specifically, Belzil and Hansen (BH) draw conclusions of the following kind: If an individual incurs certain (opportunity) costs in order to go to school, we get a lower bound on the returns to schooling, for if those returns were lower than the costs, the individual would not have chosen to go to school.[3] Conversely, every decision to exit school (or university) gives us an *upper* bound on the returns to education. One key aspect of this approach is its reliance on assumptions about the agents' intertemporal optimization behavior. For this reason, BH refer to it as "structural". From an applied perspective, the central result of their paper is that the returns to high school are much lower than the IV literature suggests. In fact, Belzil and Hansen claim that *"schooling has practically no value until grade*

[1]Popular instruments include household background variables (Rouse and Ashenfelter, 1998), institutional features of the schooling system (Angrist and Krueger, 1997), and college proximity (Card, 1995).

[2]The problem of weak instruments has been analyzed by Stock and Staiger (1997).

[3]Note that this kind of inference presumes that all individuals know how large the actual returns to schooling are.

12".

While the "revealed preference" argument sketched in the previous paragraph sounds very appealing, it entails the following problem: As we have already alluded to, the approach of BH amounts to solving – to the extent possible – a system of inequalities. The returns to a given grade of schooling should not be too large (for otherwise everybody should have completed that grade), but they should not be too small either (for otherwise everyone would have left school before that grade). Thus, the parameters of the earnings function are estimated in an attempt to rationalize all agents' schooling decisions. Accordingly, an upper bound to the returns to schooling is provided by the decisions of those students who drop out of high school before grade 12. In our view, it is not clear whether these decisions can be considered as attempts to maximize expected lifetime income.[4] Therefore, it appears prudent to take non-monetary factors into account. This, however, makes it much more difficult to measure the costs of schooling, because non-monetary variables are usually difficult to observe. In fact, BH do not even have data on the *monetary* income of an agent while in school. This raises the question of how the returns to schooling are identified in their model.

One purpose of the present paper is to address this question by examining the role of the assumption (made by BH) that the timing of an individual's labor market entry is partly driven by ex ante observable wage shocks. As it turns out, the model's identification gets lost if we abandon this assumption. Put differently, the identification strategy of BH relies on the existence of a random component in individual schooling decisions. Hence, while their approach is cast into a structural framework, it can be argued to also contain elements of the traditional IV method.

A second purpose of this paper is to examine the way in which the BH model is spec-

[4] According to Eckstein and Wolpin (1999), only 14 percent of all white male high school dropouts in the US declare to have stopped attending because they were "offered a good job". The majority (52 percent) either "did not like school", suffered from a "lack of ability, poor grades", or had been expelled.

ified. Their estimation strategy requires BH to make assumption on the distribution of unobservable characteristics in the population of agents. We analyze (computationally) how restrictive these assumptions are, i.e. whether the model is reasonably robust to distributional misspecifications. As it turns out, this is not the case. Accordingly, the estimates found by BH may suffer from a substantial misspecification bias.[5]

The organization of this chapter is as follows: In Section 3.2.1 we have a closer look at the model of Belzil and Hansen and sketch their estimation strategy. Section 3.2.2 forms the core of the paper: It contains a formal analysis of the way in which the BH model is identified. In Section 3.2.3 we examine the extent to which the estimation strategy of BH relies on distributional assumptions. Section 3.3 summarizes, discusses, and concludes.

3.2 The Approach of Belzil and Hansen

3.2.1 The Model

What we present in this section is a slightly simplified version of the empirical dynamic programming model developed by Belzil and Hansen. None of the details we have chosen to omit (for the sake of convenience) would affect any of the arguments put forward in Sections 3.2.2 and 3.2.3.[6]

---

[5] In a series of papers, Belzil and Hansen (2003, 2004, forthcoming) have applied their model to a number of additional questions. The points we make in what follows essentially apply to those other papers as well.

[6] There are three differences between the model of BH and the one we present here: First, in the model of BH there is some (exogenously given) probability that an individual who still goes to school *cannot* enter the labor market in a given period. (It enters an "interruption state" which is meant to capture events like illness.) Second, individuals in the BH model do not automatically find a job once they are on the labor market, i.e. there is (involuntary) unemployment. Third, the individuals in BH are heterogeneous with respect to some observable characteristics. These observable characteristics are assumed to be othogonal to the unobservable ones discussed below.

The model is devoted to a life cycle analysis of individual schooling decisions and has the following basic structure: Every individual enters school at the age of six and has to go there for at least six years, i.e. schooling is compulsory until the age of twelve. The maximum number of years of schooling, on the other hand, is 22. Except for these two requirements, the individuals are free to enter the labor market whenever they want to. They are assumed to "die" at the age of 65, i.e. attention is restricted to the individuals' working lives. Every individual is characterized by a two-dimensional character vector $\theta \in \mathbb{R}^2$. The first component of θ, denoted as θ^s, affects the individual's utility of going to school, the second one, denoted as θ^w, affects the individual's utility of working. The individuals are assumed to choose the timing of labor market entry so as to maximize expected lifetime utility.

Individual i's instantaneous utility of attending school during the tth year of her life can formally be described as

$$U_{it}^s = \psi(t) + \theta_i^s + \varepsilon_{it}^s, \tag{3.1}$$

The function $\psi(\cdot) : t \to \mathbb{R}$ measures the (possibly negative) "consumption value" of going to school. It has one free parameter for every (potential) year of schooling, i.e. the consumption value of schooling may vary across grades. As t is discrete and finite, we can think of $\psi(\cdot)$ as a vector of dummy variables. ε_{it}^s denotes an i.i.d. error term with mean zero and distribution function $F^s(\cdot)$. The utility of working is equal to the logarithm of earnings and can be described by the Mincerian (1974) equation

$$U_{it}^w = \ln w_{it} = \phi_1(S_i) + \phi_2 \cdot E_{it} + \phi_3 \cdot E_{it}^2 + \theta_i^w + \varepsilon_{it}^w, \tag{3.2}$$

where S_i tells us how many years of schooling individual i has completed and - just as in the case of $\psi(\cdot)$ - no structure is imposed on $\phi_1(\cdot) : S \to \mathbb{R}$, i.e. this function has as many free parameters as there are different possible values of S_i. As is standard in the human capital literature, the influence of professional experience, E, on earnings is captured by a quadratic function. ε_{it}^w is assumed to be i.i.d. with mean zero, the distribution function of ε_{it}^w is denoted as $F^w(\cdot)$. The individual characteristics θ_i^s and

θ_i^w are distributed according to a joint cumulative distribution function, denoted as $F^\theta(\cdot)$. As it will prove useful in what follows, we rewrite equation (3.1) as

$$U_{it}^s = \psi(t) + \theta_i^w + \Delta\theta_i + \varepsilon_{it}^s, \quad (3.3)$$

where $\Delta\theta_i \equiv \theta_i^s - \theta_i^w$ denotes individual $i's$ "comparative advantage" at school.

The lifetime utility of individual i is given by $U_i = \sum_{t=1}^{T} \delta^t \cdot U_{it}$ where $T = 53$, i.e. we neglect the individual's first twelve years of life (because the utility experienced during those years is exogenously given). Since no agent can ever return to school once she has entered the labor market and S_i (the number of years of schooling) can only take on a relatively small number of discrete values, we can easily compute the individual's expected lifetime utility for all possible values of S_i. The problem of maximizing U_i, however, is complicated by the fact that the realizations of ε_{it}^s and ε_{it}^w, denoted as ϵ_{it}^w and ϵ_{it}^s, are assumed to be observable prior to period t, i.e. the individual can condition her entry decision on these shocks. For this reason, the agents cannot decide ex-ante for how many years they will go to school, but they have to solve the maximization problem sequentially.

Individual i decides to enter the labor market in period t if and only if $\epsilon_{it}^w - \epsilon_{it}^s > V_{it}^s - V_{it}^w(S_i = t-1)$, where V_{it}^s and $V_{it}^w(S_i = t-1)$ denote the expected values of going to school and entering the labor market after $t-1$ years of schooling, respectively. $V_{it}^w(S_i)$ is recursively given by

$$V_{it}^w(S_i) = \phi_1(S_i) + \phi_2 \cdot E_{it} + \phi_3 \cdot E_{it}^2 + \theta_i^w + \delta \cdot V_{it+1}^w(S_i) \quad (3.4)$$

Since we know that $V_{iT+1}^w(\cdot) = 0 \; \forall \; S_i$, $V_{it}^w(\cdot)$ can easily be computed for all values of t and S_i by means of backward induction. The value of remaining in school can be expressed by the Bellman equation

$$V_{it}^s = \psi(t) + \theta_i^w + \Delta\theta_i + \delta \cdot E\left[Max\{V_{it+1}^s + \varepsilon_{it}^s, V_{it+1}^w(S_i = t) + \varepsilon_{it}^w\}\right] \quad (3.5)$$

Note that the last term on the right hand side of equation (3.5) does not simplify to $\delta \cdot Max\{V_{it+1}^s, V_{it+1}^w(S_i = t)\}$ (even though both ε_{it}^s and ε_{it}^w have mean zero), because the variance of ε_{it}^s and ε_{it}^w brings about some option value. At this point we can already make the following useful observation (which is new to our paper):

Lemma 5: *The difference* $V_{it}^s - V_{it}^w(S_i = t-1)$ *only depends on* θ^w *and* θ^s *through* $\Delta\theta_i$. *Moreover,* $V_{it}^s - V_{it}^w(S_i = t-1)$ *is a strictly increasing function of* $\Delta\theta_i$.

Proof. The proof is by induction. We first show that the lemma is true for every agent's final year of potential schooling, i.e. for $t = 16$. (Recall that we do not consider the first six years of schooling, because those are compulsory.) We then go on to show that, if the lemma is true for any t, then it must also be true for $t-1$. As for the first step, note that all agents have to leave school at the age of 28, so equation (3.5) simplifies to $V_{i16}^s = \psi(16) + \theta_i^w + \Delta\theta_i + \delta \cdot V_{i17}^w(16)$. Taking the difference of this term and $V_{i16}^w(S_i = 15)$ we directly see that the result depends on θ^w and θ^s only through $\Delta\theta_i$ and is strictly increasing in $\Delta\theta_i$. Consider next the following reformulation of equation (3.5): $V_{it}^s = \psi(t) + \theta_i^w + \Delta\theta_i + \delta \cdot \left(V_{it+1}^w(S_i = t) + Max\{V_{it+1}^s - V_{it+1}^w(S_i = t) + \varepsilon_{it}^s; \varepsilon_{it}^w\}\right)$. Presupposing that the lemma holds for $V_{it+1}^s - V_{it+1}^w(S_i = t)$, the same must - once again by simple inspection - be true for $V_{it}^s - V_{it}^w(S_i = t-1)$. ■

Intuitively, Lemma 5 must be true, because θ^w enters linearly into both (3.2) and (3.3), whereas θ^s does not appear in either of the two equations (except through $\Delta\theta_i$). Hence, no matter what agent i does, she will "get" θ_i^w in every period of her life and θ_i^s in none of them.

Lemma 5 has an important implication: Since individual i enters the labor market in period t if and only if $\epsilon_{it}^w - \epsilon_{it}^s > V_{it}^s - V_{it}^w(S_i = t-1)$, the probability of an individual leaving school in any given period (conditional on not having left it before) is decreasing in $\Delta\theta$. Hence, (average) schooling attainments must be increasing in $\Delta\theta$. In order to check whether this conclusion is in line with the results of BH, let us have a look at their estimates of the distribution of unobserved abilities:

	θ^s	θ^w	$\Delta\theta$
type 1	$\theta_1^s = -0.732$	$\theta_1^w = 2.140$	$\Delta\theta_1 = -2.871$
type 2	$\theta_2^s = -1.102$	$\theta_2^w = 1.679$	$\Delta\theta_2 = -2.781$
type 3	$\theta_3^s = -0.879$	$\theta_3^w = 1.914$	$\Delta\theta_3 = -2.792$
type 4	$\theta_4^s = -1.321$	$\theta_4^w = 1.377$	$\Delta\theta_4 = -2.698$
type 5	$\theta_5^s = -1.182$	$\theta_5^w = 1.549$	$\Delta\theta_5 = -2.730$
type 6	$\theta_6^s = -1.490$	$\theta_6^w = 1.082$	$\Delta\theta_6 = -2.572$

Table 1: The Distribution of Abilities in Belzil and Hansen

The first two columns of Table 1 are directly taken from Table III in BH, the third one is based on own computations. As apparent from columns 2 and 3 of the table, BH find that more able workers have a comparative *disadvantage* at school. Hence, we would - by Lemma 1 - expect a negative correlation between market ability and schooling. Yet, BH find that correlation to be significantly positive (0.26). This contradiction indicates that, unless Lemma 5 is faulty, something must be wrong with the estimates of BH.

BH use maximum likelihood techniques in order to jointly estimate $\psi(\cdot)$, $\phi_1(\cdot)$, ϕ_2, ϕ_3, δ, $F^\theta(\cdot)$, $F^w(\cdot)$, and $F^s(\cdot)$. This requires them to make assumptions on the functional forms of $F^\theta(\cdot,\cdot)$, $F^w(\cdot)$, and $F^s(\cdot)$. They take ε_{it}^w and ε_{it}^s to be normally distributed with mean zero, whereas θ is assumed to follow a discrete distribution with six mass points (Heckman and Singer, 1984). Thus, while $F^w(\cdot)$, and $F^s(\cdot)$ have just one free parameter each, $F^\theta(\cdot)$ is a very flexible function: 12 parameters are needed in order to characterize the six two-dimensional values that θ can assume (θ^1,..., θ^6). In addition, we need to know the probabilities with which these values occur, denoted as p^1,..., p^6. As these probabilities need to add up to one, five additional degrees of arise.

BH only have data on U^w (wages) and S (schooling attainments). Hence, the likelihood function for a given agent, L_i, consists of three parts: The probability of remaining at school for (at least) S_i years, the probability of entering the labor market after S_i years of schooling and the likelihood of matching the wage profile w_i in all subsequent years. Since θ can assume six different values, we need to compute the likelihood function for all types separately. The complete likelihood function is found by taking the product of the three parts, weigting it by p^k, and then summing up over all k and i.

3.2.2 Identification

BH claim that "*identification of the wage return to schooling [...] is relatively straightforward given panel data on wages [...] and, hence, does not require discussion.*" By contrast, we do not find the way in which the BH model is identified to be straightforward. Therefore, the present section is devoted to an analysis of the conditions under which the identification result holds. More specifically, we focus on the role of the assumption that ϵ_{it}^w and ϵ_{it}^s (the realizations of ε_{it}^w and ε_{it}^s) are observable prior to period t. The following proposition shows that, if we relax this assumption, the model is not identified anymore.

Proposition 13: *If* ϵ_{it}^w and ϵ_{it}^s *were not observable prior to period* t*, the model of BH would not be identified.*

Proof. The proof consists of showing that, if ϵ_{it}^w and ϵ_{it}^s were not observable, different sets of parameters could generate exactly the same distribution of observable variables. For simplicity, we only consider (joint) changes in $\psi(\cdot)$, $\phi_1(\cdot)$, and $F^\theta(\theta^w|\Delta\theta)$, taking ϕ_2, ϕ_3, δ, $F^w(\cdot)$, $F^s(\cdot)$, and $F^{\Delta\theta}(\cdot)$ as fixed. Since $V_{it}^s - V_{it}^w$ is - by Lemma 5 - a strictly increasing function of $\Delta\theta_i$, this simplification implies that we can take the order of schooling attainments as given. Our argument consists of two parts: We first demonstrate that, no matter what $\phi_1(\cdot)$, and $F^\theta(\theta^w|\Delta\theta)$ look like, we can generate any arbitrary distribution of schooling attainments (which respects the order prescribed by $F^{\Delta\theta}(\cdot)$) by means of $\psi(\cdot)$ alone. We then go on to show that, holding the distribution of schooling attainments fixed, there are different combinations of $\phi_1(\cdot)$ and $F^\theta(\theta^w|\Delta\theta)$ which yield precisely the same wage profile w. Taken together, the two steps imply that we cannot infer the set of true parameter values from data on S and w.

The proof of the argument's first part is by construction: Note that we can characterize the distribution of schooling attainments by a set of threshold values with respect to $\Delta\theta$, denoted as τ. All agents with $\Delta\theta > \tau_{16}$ go to school for 16 years, those with $\tau_{16} \geq \Delta\theta > \tau_{15}$ go there for 15 years etc. Students with $\tau_1 \geq \Delta\theta$ complete only the minimum amount of schooling. In order to find the function $\psi(\cdot)$ which generates a given distribution τ, we need to do the following: We substitute τ_1 for $\Delta\theta$ in equation (3.5), take the difference $V^s_{i16} - V^w_{i16}$, set it to zero, and then solve for $\psi(16)$. Once $\psi(16)$ is known, we can use the same solution method to determine $\psi(15)$ etc.

As for the proof of the argument's second part, recall that $\ln w_{it} = \phi_1(S_i) + \phi_2 \cdot E_{it} + \phi_3 \cdot E^2_{it} + \theta^w_i + \varepsilon^w_{it}$ (where ϕ_2, ϕ_3, and the distribution of ε^w_{it} are all taken as given). We want to show that different combinations of $\phi_1(\cdot)$ and $F^\theta(\theta^w|\Delta\theta)$ generate identical wage patterns. In doing so, we do not have to take the effect of changes in $\phi_1(\cdot)$ and θ^w on schooling attainments into account (as we have just seen that these effects can always be compensated for by changes in $\psi(\cdot)$). Thus, $\phi_1(\cdot)$ and $F^\theta(\theta^w|\Delta\theta)$ are not separately identified, as any increase in $\phi_1(t)$ can be offset by a decrease in θ^w for all agents with $\tau_{t+1} \geq \Delta\theta > \tau_t$. ■

Proposition 13 can be strengthened in two respects: First, as long as every level of schooling attainment is chosen by *some* individuals, the non-identifiability of the modified model can also be proven by considering joint changes in $\phi_1(\cdot)$ and $F^\theta(\cdot,\cdot)$. Second, ϵ^w_{it} and ϵ^s_{it} do not necessarily have to be unobservable prior to period t. It is enough for them not to have any influence on the individuals' schooling decisions.

In what way does the observability of ϵ^w_{it} and ϵ^s_{it} help us to identify the returns to schooling? The answer to this question is most easily understood by bringing to one's mind the way in which observable wage shocks influence the timing of labor market entry: If there is a positive shock, every individual will - ceteris paribus - tend to enter the labor market, otherwise she will prefer to remain in school. Hence, if an individual leaves school even though she is subject to a negative wage shock, that individual probably has spent some time waiting in vain for a good wage offer. Conversely, if an individual enters the labor market in response to a large positive shock, that individual may well have left school earlier than expected. Thus, there should - for a given level of schooling - be a positive correlation between the initial wage shock experienced by an individual and her comparative advantage at school, $\Delta\theta$. This implies that the random part of an individual's starting salary can play a role comparable to that of a standard instrument in reduced-form models: By bringing about some exogenous variation in schooling attainments, the observability of ϵ_{it}^{w} and ϵ_{it}^{s} allows us to compare individuals with identical characteristics, but different schooling attainments.

One natural question to ask at this point is whether wage shocks can legitimately be used as a "quasi-instrument" for schooling attainments. In order for the answer to this question to be "yes", two conditions need to be satisfied: First, individual labor market entry decisions should be related to contemporaneous wage offers. Second, the nature of this relationship should be independent of an individual's type. While we believe that both of these conditions can be checked, it is beyond the scope of this paper to carry out such an analysis.

3.2.3 Specification

Our focus in this section will be on the assumption that the distribution of unobserved abilities consists of six mass points. At first glance, this assumption appears to be relatively weak: After all, $F^{\theta}(\cdot)$ has (as we have already noted in Section 3.2.1) no less than 17 degrees of freedom. Nevertheless, the true distribution of unobserved abilities is potentially continuous, and it is not easy to cover the whole of $\mathbb{R}^2$ by just six points. The question is how the model rationalizes the choices of those agents whose actual characteristics are far from any of the estimated mass points. The answer to this question is simple: These individuals must have experienced very substantial shocks. Such shocks can only occur if the variances of ε_{it}^{w} and ε_{it}^{s}, denoted as σ_w and σ_s in the following, are sufficiently large. Since the evolution of wages over time provides direct evidence on the variability of the wage shocks, σ_w cannot "pick up the burden". Therefore, misspecifications of $F^{\theta}(\cdot)$ are bound to results in an overestimation of σ_s. (According to the estimates of BH, it is more than five times as large as σ_w.) As we have already alluded to in Section 3.2.2, this has important consequences for the agents' incentives to remain in school. Let us have a closer look at the relationship between σ_w and V_{it}^{s}. From equation (3.5) we know that

$$V_{it-1}^{s} = E\left[Max\{V_{it+1}^{s} + \varepsilon_{it}^{s}, V_{it+1}^{w} + \varepsilon_{it}^{w}\}\right] \tag{3.6}$$

where, for convenience, we have set $\psi(t) + \theta_i^w + \Delta\theta_i$ to zero and δ to one. Dropping, in addition, the subindex i, we can rewrite this equation as

$$V_{t-1}^s = E\left[V_t^w + \varepsilon_{t-1}^w + Max\{V_t^s - V_t^w + \varepsilon_{t-1}^s - \varepsilon_{t-1}^w, 0\}\right]$$

$$= V_t^w + E\left[Max\{V_t^s - V_t^w + \varepsilon_{t-1}^s - \varepsilon_{t-1}^w, 0\}\right]$$

$$= V_t^w + \frac{E\left[V_t^s - V_t^w + \varepsilon_{t-1}^s - \varepsilon_{t-1}^w | \varepsilon_{t-1}^s - \varepsilon_{t-1}^w > V_t^w - V_t^s\right]}{P\left(\varepsilon_{t-1}^s - \varepsilon_{t-1}^w > V_t^w - V_t^s\right)}$$

$$= V_t^w + \Phi\left(\frac{V_t^s - V_t^w}{\sigma}\right) \cdot (V_t^s - V_t^w) + \int_{V_t^s - V_t^w}^{\infty} \frac{x \cdot \exp(-x^2/2\sigma^2)}{\sigma \cdot \sqrt{2\pi}} dx$$

$$= V_t^w + \Phi\left(\frac{V_t^s - V_t^w}{\sigma}\right) \cdot (V_t^s - V_t^w) + \frac{\sigma}{\sqrt{2\pi}} \cdot \exp\left(-\frac{(V_t^s - V_t^w)^2}{2\sigma^2}\right) \qquad (3.7)$$

where $\sigma = \sqrt{\sigma_w^2 + \sigma_s^2}$. The option value brought about by a large variance of ε_{it}^w is contained in the last term on the right hand side of (3.7): Both $\sigma/\sqrt{2\pi}$ and $\exp(-\left(V_t^s - V_t^w\right)^2/2\sigma^2)$ are increasing in σ.

Intuitively, the positive effect of σ on V^s can be understood as follows: As long as an agent goes to school, she can, in any given period, choose whether or not to enter the labor market. This gives her the opportunity to avoid large negative shocks or take advantage of positive ones, i.e. it generates a positive option value of going to school. The size of this effect depends on σ, because substantial shocks are more likely to occur if σ is large. We now come to the point of understanding the link between misspecifications of $F^\theta(\cdot)$ and the estimation of the earnings function: Since both σ and the returns to schooling provide an incentive not to leave school too early, they are substitutes in explaining agents' schooling decisions. Therefore, if σ gets overestimated, the converse must be true of the returns to schooling, for otherwise the choices of those students who drop out of high school early on could not be rationalized.

In order to get an idea of how large this effect is, we apply the BH estimator to two different sets of artificially generated data. The first data set is well behaved, i.e. it is generated under the same assumption on which the BH estimator relies. By contrast, the second data set takes the distribution of agents' characteristics to be continuous. Thus, we test whether the BH estimator is able to uncover the true returns to schooling when the assumption of a discrete type distribution fails, using the first data set as a control.

The model we consider is a slightly simplified version of the one presented in Section 3.2.1: In order to alleviate the computational burden, every individual is assumed to live for only 10 periods of which she can spend at most the first three ones going to school. In the first data set, the distribution of agents' characteristics consists of four mass points. The assumption that there are only four different values of $\Delta\theta$ also holds true in the second data set. Given any of these four values, however, agents' labor market ability is normally distributed with mean μ_θ^k and standard variation σ_θ^k. Our estimation procedure (which is the same for both data sets) is based on the assumption that there are four different types of agents and, for simplicity, takes the consumption value of schooling, the returns to experience, the discount rate, and agents' comparative advantage at school to be known. Thus, we only estimate the returns to schooling, the standard deviations σ_w and σ_s, and the distribution of agents' market ability.

Our estimates can be found in Table 2. The first column ("estimate 1") corresponds to the well-behaved data set while the second column ("estimate 2") corresponds to data set which is based on the assumption that the distribution of agents' types is continuous.[7]

[7]The program used to generate these results was written in MATLAB and is available from the author upon request.

Parameter	Value	Estimate 1	Estimate 2
$\phi_1(1)$	0.1	0.099 (0.002)	0.048 (0.008)
$\phi_1(2)$	0.2	0.197 (0.003)	0.135 (0.011)
$\phi_1(3)$	0.3	0.298 (0.002)	0.289 (0.010)
σ_w	0.5	0.499 (0.002)	0.682 (0.009)
σ_s	0.5	0.514 (0.025)	1.130 (0.039)
θ_1^w	–	1.001 (0.015)	0.402 (0.083)
θ_2^w	–	1.987 (0.013)	1.843 (0.067)
θ_3^w	–	3.011 (0.008)	3.219 (0.070)
θ_4^w	–	4.000 (0.011)	4.714 (0.085)

Table 2: Results of the Computational Experiments

Parameter values: $\psi(1) = \psi(2) = \psi(3) = 0$, $\delta = 1$, $\phi_2 = \phi_3 = 0$, $\mu_\theta^1 = -2$, $\mu_\theta^2 = -1.4$, $\mu_\theta^3 = -0.8$, $\mu_\theta^4 = -0.2$, $\sigma_\theta^k = 0\ \forall k$ (first data set), $\sigma_\theta^k = 1\ \forall\ k$ (second data set). Number of observations: $n = 1000$. Standard errors are given in parentheses.

The results which are contained in Table 2 confirm our theoretical predictions. If the data generating process "respects" the assumptions of the BH estimator, all parameters are estimated at high precision. Misspecifications of the distribution of agents' characteristics, however, result in an overestimation of σ which, in turn, brings about downwardly biased estimates of the returns to schooling. Note that the returns to the final year of schooling are largely unaffected by this problem. Again, this is in line with our theoretical considerations. The final year of schooling cannot contain any option value, because all agents *have* to enter the labor market afterwards. Overall, our estimates look very much like the ones of Belzil and Hansen: The returns to schooling are very low until "grade 2", and the earnings function is strongly convex in schooling. This, of course, does not automatically imply that the results of Belzil and Hansen are biased. It *does* suggest, however, that the possibility of a specifications bias needs to be taken seriously.

3.3 Conclusion

The purpose of this paper has been to critically evaluate the "structural" approach to estimating the returns to schooling as advocated by Belzil and Hansen. We have first shown that the identification of the BH model hinges on the assumption that individual schooling decisions are driven by observable random shocks. If this assumption fails, the model is not identified. We have then gone on to demonstrate - both theoretically and computationally - that the way in which BH have specified the distribution function of agents' characteristics is less flexible than one might think. Accordingly, the model reacts sensitively to distributional misspecifications. The resulting specification biases are able to account for the very low estimates of the returns to schooling found by BH.

The criticism raised in this paper should not be interpreted as a criticism of structural estimation in general. It undoubtedly makes sense to estimate the parameters of an economic model rather than simply positing some – usually linear – causal relationship. Yet, while (micro-)economic models have the virtue of being explicit and coherent, they cannot bring about any "identification miracles". Technically, sufficiently complex models may allow us to identify parameters which are not identifiable in reduced-form models, but these identification strategies are unlikely to be reliable. Accordingly, our answer to the question asked in the title of this paper reads as follows: Yes, we can trust structural estimates of the returns to schooling, but we should not trust identification strategies which are not amenable to economic intuition. Hence, while structural estimates may help us to refine the insights produced by reduced-form models, they are unlikely to bring about an econometric revolution.

More radical progress may be achievable if we modify the standard approach to estimating the earnings function at a different end. Both Belzil and Hansen and most of the IV literature make two steps at once by directly considering the relationship between schooling attainments and subsequent earnings. In our view it would be more reasonable to first examine to what extent (and in what way) certain (cognitive, social, or technical) skills are acquired at school and to then analyze the value of these skills in a given job. (Clearly, different jobs require different skills, so the returns to schooling will generally depend on what kind of job we look at.) Not only do we think that the answers to these two questions are easier to obtain than those sought after by the existing literature. If we manage to find them, they will also provide us with a much more complete picture of how our education system works and how we can improve it.

Bibliography

[1] Alesina, Alberto, Rafael di Tella, and Robert McCulloch (2004): "Inequality and Happiness: Are Americans and Europeans Different?" *Journal of Public Economics* 88, 2009-2042

[2] Angrist, Joshua, and Alan Krueger (1991): "Does Compulsory School Attendance Affect Schooling and Earnings", *Quarterly Journal of Economics* 106, 75-102

[3] Athey, Susan and Philip Haile (2002): Identification of Standard Auction Models, *Econometrica* 70, 2107-2140

[4] Bajari, Patrick and Ali Hortacsu (2003): The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions, *Rand Journal of Economics* 3, 329-355

[5] Belzil, Christian, and Jörgen Hansen (2002): "Unobserved Ability and the Return to Schooling", *Econometrica* 70, 2075-2091

[6] Belzil, Christian, and Jörgen Hansen (2003): "Structural Estimates of the Intergenerational Education Correlation", *Journal of Applied Econometrics*, 679-699

[7] Belzil, Christian, and Jörgen Hansen (2004): "Earnings Dispersion, Risk Aversion, and Education", *Research in Labor Economics* 335-358

[8] Belzil, Christian, and Jörgen Hansen (forthcoming): "A Structural Analysis of the Correlated Random Coefficient Wage Regression Model", *Journal of Econometrics*

[9] Card, David (1995): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling", in: *Aspects of Labor Market Behavior: Essays in Honor of John Vanderkamp* (Christofides, Grant, and Swidinsky, ed.), Toronto: University of Toronto Press

[10] Card, David (1999): "The Causal Effect of Schooling on Earnings", in: *Handbook of Labor Economics 3A* (Ashenfelter and Card, eds.), Amsterdam: Elsevier Science / North Holland

[11] Cho, Inn-Koo and David Kreps (1987): "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics* 102, 179-221

[12] Clark, Andrew, Fabrice Etilé, Fabien Postel-Vinay, Claudia Senik, and Karine van der Straeten (2005): "Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries." *Economic Journal* 115, C118-C132

[13] Cole, Harold, George Mailath, and Andrew Postlewaite (1992): "Social Norms, Saving Behavior, and Growth." *Journal of Political Economy* 100, 1092-1125

[14] Corneo, Giacomo and Hans-Peter Grüner (2000): "Social Limits to Redistribution." *American Economic Review* 90, 1491-1507

[15] Diamond, Peter (1971): A Model of Price Adjustment, *Journal of Economic Theory* 3, 156-168

[16] Eckstein, Zvi, and Kenneth Wolprin (1999): "Youth Unemployment and Academic Performance in High School", *Econometrica* 67, 1295-1339

[17] Frey, Bruno and Alois Stutzer (2002): "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature* 40, 402-435

[18] Fuente, Angel de la (2000): *Mathematical Methods and Models for Economists*, Cambridge University Press

[19] Gale, Douglas (1987): Limit Theorems for Markets with Sequential Bargaining, *Journal of Economic Theory* 43, 20-54

[20] Gale, David and Lloyd Shapley (1962): "College Admission and the Stability of Marriage." *American Mathematical Monthly* 69, 9-15

[21] Griliches, Zvi (1970): "Notes on the Role of Education in Production Functions and Growth Accounting" in: *Education, Income and Human Capital* (Lee Hansen, ed.), New York: Columbia University Press

[22] Griliches, Zvi (1977): "Estimating the Returns to Schooling: Some Econometric Problems" *Econometrica* 45, l-22

[23] Haile, Philip (2003): Auctions with Private Uncertainty and Resale Opportunities, *Journal of Economic Theory* 108, 72-110

[24] Heckman, James, and Burton Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", *Econometrica* 52, 271-320

[25] Hirsch, Fred (1976): *Social Limits to Growth*, Harvard University Press

[26] Hopkins, Edward and Tatiana Kornienko (2004): "Running to keep in the same Place: Consumer Choice as a Game of Status." *American Economic Review* 94, 1085-1107

[27] Katz, Lawrence and David Autor (1999): "Changes in the Wage Structure and Earnings Inequality." *Handbook of Labor Economics* 3A, North-Holland

[28] Keane, Michael, and Kenneth Wolpin (1997): "The Career Decisions of Young Men", *Journal of Political Economy* 105, 473-522

[29] Krishna, Vijay (2002): *Auction Theory*, Academic Press

[30] Lauermann, Stephan (2006): Dynamic Matching and Bargaining Games: A General Approach, *Working Paper*

[31] Lauermann, Stephan and Wolfram Merzyn (2006): "Failures of the Monotone Likelihood Ratio Property in Auctions with a Stochastic Number of Bidders", *Working Paper*

[32] Mailath, George (1987): "Incentive Compatibility in Signaling Games with a Continuum of Types." *Econometrica* 55, 1349-1365

[33] Mas-Colell, Andreu, Michael Whinston, and Jerry Green (1995): *Microeconomic Theory*, Cambridge University Press

[34] Mezzetti, Claudio, Aleksandar Pekec and Ilia Tsetlin (2004): Sequential vs. Single-Round Uniform-Price Auction, *Working Paper*

[35] Milgrom, Paul (1981): Rational Expectations, Information Acquisition, and Competitive Bidding, *Econometrica* 49, 921-943

[36] Milgrom, Paul and Robert Weber (1982): A Theory of Auctions and Competitive Bidding, *Econometrica* 50, 1089-1122

[37] Milgrom, Paul and Robert Weber (2000): A Theory of Auctions and Competitive Bidding II, in: *The Economic Theory of Auctions (P. Klemperer, ed.)*, Edward Elgar

[38] Milgrom, Paul and Ilya Segal: Envelope Theorems for Arbitrary Choice Sets, *Econometrica* 70, 583-601

[39] Mincer, Jacob (1974): "*Schooling, Experience, and Earnings*", New York: National Bureau of Economic Research

[40] Mood, Alexander, Franklin Graybill, and Duane Boes (1974): *Introduction to the Theory of Statistics*, McGraw-Hill

[41] Pesendorfer, Wolfgang and Jeroen Swinkels (1997): The Loser's Curse and Information Aggregation in Common Value Auctions, *Econometrica* 65, 1247-1282

[42] Pesendorfer, Wolfgang and Jeroen Swinkels (2000): Efficiency and Information Aggregation in Auctions, *American Economic Review* 90, 499-525

[43] Postlewaite, Andrew (1998): "The Social Basis of Interdependent Preferences." *European Economic Review* 42, 779-800

[44] Riley, John (1979): "Informational Equilibrium." *Econometrica* 47, 331-359

[45] Rochet, Jean-Charles and Lars Stole (2003): "The Economics of Multidimensional Screening." *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress,* Cambridge University Press

[46] Rockafellar, Ralph (1970): *Convex Analysis*, Princeton University Press

[47] Rouse, Cecilia, and Orley Ashenfelter (1998): "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins", *Quarterly Journal of Economics* 113, 253-284

[48] Satterthwaite, Mark A. and Artyom Shneyerov (forthcoming): Dynamic Matching, Two-sided Incomplete Information, and Participation Costs: Existence and Convergence to Perfect Competition, *Econometrica*

[49] Schor, Juliet (1992): "*The Overworked American: The Unexpected Decline of Leisure*", Basic Books

[50] Serrano, Roberto and O. Yosha (1995): Decentralized Markets with Pairwise Meetings: Recent Developments, *Journal of International and Comparative Economics 4*, 223-241

[51] Simon, Carl and Lawrence Blume (1994): *Mathematical Methods for Economists*, Norton & Company

[52] Staiger, Douglas, and James Stock (1997): "Instrumental Variables Regression with Weak Instruments", *Econometrica* 65, 557-586

[53] Wilson, Robert (1969): A Bidding Model of Perfect Competition, *Review of Economic Studies 4,* 511-518

[54] Wolinsky, Asher (1990): Information Revelation in a Market with Pairwise Meetings, *Econometrica 58,* 1-23