

Kemptoner, Daniel; Tolan, Songül

**Working Paper**

## The role of time preferences in educational decision making

DIW Discussion Papers, No. 1628

**Provided in Cooperation with:**

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* Kemptoner, Daniel; Tolan, Songül (2016) : The role of time preferences in educational decision making, DIW Discussion Papers, No. 1628, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/148562>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# 1628

## Discussion Papers

Deutsches Institut für Wirtschaftsforschung

2016

# The Role of Time Preferences in Educational Decision Making

Daniel Kemptner and Songül Tolan

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

#### IMPRESSUM

© DIW Berlin, 2016

DIW Berlin  
German Institute for Economic Research  
Mohrenstr. 58  
10117 Berlin

Tel. +49 (30) 897 89-0  
Fax +49 (30) 897 89-200  
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:  
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:  
<http://ideas.repec.org/s/diw/diwwpp.html>  
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

# The Role of Time Preferences in Educational Decision Making\*

Daniel Kemptner<sup>a</sup> and Songül Tolan<sup>a,b</sup>

<sup>a</sup>Deutsches Institut für Wirtschaftsforschung, Berlin

<sup>b</sup>Free University of Berlin

September 2016

## Abstract

We analyze the implication of time-inconsistent preferences in educational decision making and corresponding policies using a structural dynamic choice model. Based on a novel identification approach, we exploit variation in average years invested in degree attainment through various educational reforms to identify the discount factor of hyperbolic time preferences. We make two important research contributions. First, we estimate our model using data from the German Socioeconomic Panel (SOEP) and provide quantitative evidence for time-inconsistent behavior in educational decision making. Second, we evaluate the relevance of time-inconsistent behavior for the effectiveness of education policies. For this purpose, we simulate policies where time preferences may play an important role: (1) an increase in the state grant for students as a way to affect short-term costs while at school and (2) an increase in the state grant as a loan that must be paid back after education is completed. We find substantial differences in the educational outcomes when comparing them to the outcomes based on a model specification with exponential discounting. Hence, the common assumption of exponential discounting in educational decisions may be too restrictive.

**JEL:**C61,D91, I21

**Keywords:** time preferences, hyperbolic discounting, structural estimation, dynamic discrete choice, education

---

\*Valuable comments by Peter Haan, Jaap Abbring, and Ulrich Schneider as well as by participants at the EEA-ESEM Congress 2016, and the Applied Microeconomics Seminars at Tilburg University and the DIW are gratefully acknowledged. Daniel Kemptner would like to thank Thyssen Foundation (project: 10141098) for financial support for this project.

# 1 Introduction

Economists have long understood the key role of education for economic growth (Barro, 1998; Lindahl and Krueger, 2001). The estimates of the returns to one additional year of education range between 6 and 16% (Card, 1999). As a consequence, many educational policies aim at increasing educational investments of students (e.g. student grants or loans). While the design of successful education policies requires a good understanding of the underlying mechanisms of educational choices, the intertemporal preferences of students are not yet fully understood. In this study, we contribute to closing this gap in our knowledge by investigating the role of time discounting within a structural dynamic choice model.

The famous Stanford marshmallow experiment first highlighted the strong positive link between patience and educational attainment (Mischel et al., 1989). We extend this literature by analyzing how much the functional form of patience in an economic choice model matters for policy evaluation. We do this by deviating from the assumption of exponential discounting and allow for time-inconsistent preferences through hyperbolic discounting. In the spirit of Magnac and Thesmar (2002), Fang and Wang (2015) and Chan (2013), we use a novel identification method to not only identify an exponential discount factor, but also an additional parameter that captures hyperbolic discounting.<sup>1</sup> Behavioral responses to education policies hinge on intertemporal preferences because individuals trade off short-term costs against potential future returns on the educational investments. Hence, the way people discount is likely to have an important impact on the effectiveness of policies that decrease the short-term cost of education or increase its long-term benefits.

In this paper, we make two important research contributions. First, we use the German Socioeconomic Panel (SOEP), to estimate a dynamic structural model of educational choices that allows for hyperbolic discounting. The estimation is based on a sample of West German students. In line with Magnac and Thesmar (2002), Chan (2013), Fang and Wang (2015), and Haan et al. (2016) we achieve identification by imposing exclusion restrictions that affect educational choices indirectly through their impact on the transition probabilities of relevant state variables but have no impact on current utility flows. These restrictions reveal information on the discounting behavior of the observed individuals. We use birth cohort groups as well as regions

---

<sup>1</sup>The analysis of potential deviations from standard assumptions in economic models is receiving increasing interest in the empirical literature (Fang and Silverman, 2006). One formalization of time-inconsistent preferences is hyperbolic discounting (Laibson, 1997; O'Donoghue and Rabin, 1999; O'Donoghue and Rabin, 1999). While the hyperbolic discounter discounts exponentially between any two subsequent payoffs made in a distant future, she might change her preferences and put a higher weight on the more immediate payoff once she approximates the point in time when the first payoff is made.

that were affected by different educational policy reforms as exclusion restrictions and show their relevance on the time invested for the attainment of educational degrees. Agents are assumed to face two kinds of uncertainty: (1) there is uncertainty over whether an additional year invested into education will, in fact, be successful and lead to a degree (affected by the policy reforms); and (2) there is uncertainty over the returns to the degree earned when exiting education. The estimation of the structural parameters of the choice model indicates time-inconsistent behavior and provides quantitative evidence for its relevance.

Our second contribution to the literature refers to an evaluation of the relevance of time-inconsistent behavior for the effectiveness of education policies. For this purpose, we simulate two policies where time preferences may play an important role: (1) an increase in the state grant for students as a way to affect short-term costs while at school; and (2) an increase in the state grant as a loan which will have to be paid back after the end of the education. We investigate whether present-biased preferences matter for educational decisions depending on whether financial support during the educational time is a grant or a loan. We find substantial differences in the effects of these policies when comparing educational outcomes based on a model specification with hyperbolic discounting with the ones based on a specification with exponential discounting. Furthermore, the response to the two policies differs more for exponential than for hyperbolic discounters.

Aiming at a better understanding of educational decisions, other studies analyze the role of socio-economic background (Cameron and Heckman, 2001), preferences and abilities (Eckstein and Wolpin, 1999), uncertainty and sequential information updates (Carneiro et al., 2003; Heckman et al., 2005), as well as individuals' risk aversion (Belzil and Leonardi, 2007).

Experimental results in behavioral economics indicate hyperbolic behavior in intertemporal decision making (see Giné et al. (2010) for a review and Cohen et al. (2016) for a discussion of the predictive power of such experiments). There are a number of studies implementing quasi-hyperbolic discounting in dynamic discrete choice models – an approximation to hyperbolic discounting in discrete time (Laibson, 1997). Magnac and Thesmar (2002) show that basic dynamic structural models are underidentified if the discount factor is estimated along with the other structural parameters. A large share of the literature addresses this identification problem and tries to estimate the discount factor in both exponential and hyperbolic settings. For instance, Fang and Silverman (2009) estimate the discount factor in a dynamic choice model of labor supply and welfare take-up of single mothers. Their results suggest a significant present bias factor and a better fit to the data when allowing for hyperbolic discounting. Similarly, Paserman (2008) rejects the hypothesis

of exponential discounting for low-wage workers when implementing and estimating hyperbolic discounting in a job search model. Further literature on the implementation and/or estimation of hyperbolic discounting include Laibson et al. (2007), who estimate short and long term discount rates in a structural buffer-stock consumption model, DellaVigna and Paserman (2005), who study the role of time preferences in job search, Tobacman and Skiba (2005), who explain the behavior of payday loan borrowers with hyperbolic discounting, and Gustman and Steinmeier (2012), who looks at hyperbolic discounting in the context of retirement.

In our study, we exploit two advantages of structural models: (1) the possibility to investigate channels that may drive the schooling decision; and (2) the possibility to simulate counterfactual policy changes. For this purpose, we set up a life-cycle model in a dynamic discrete choice framework. In education economics, this approach is pioneered by Keane and Wolpin (1997) and discussed by Heckman et al. (2016). Keane and Wolpin (1997) use their parameter estimates to simulate the effect of a college fee subsidy on educational decisions while risk attitudes or time preferences remain unidentified. Our study relies on the basics of the model formulated by Belzil and Hansen (2002). It is an optimal stopping model in which the agents make annual decisions to remain in education or to exit to the labor market. Our work relates to Belzil and Hansen (1999) and Oosterbeek and van Ophem (2000), who study the role of the discount factor in education and its relation to the socioeconomic background based on a model with exponential discounting.

This paper is structured as follows. First, we describe the institutional background and educational reforms that are relevant to our identification strategy. Then, we develop the model. Identification, model solution, and estimation results are discussed in the subsequent sections. In section 7, we show the estimation results and model fit. Section 8 presents the policy simulations and section 9 concludes.

## 2 Institutional Background

Compulsory elementary school starts at age six and typically lasts four years until the age of ten. Subsequently, students continue their education on one of three tracks: (1) lower secondary education (five years); (2) intermediate secondary education (six years; typically preparing for vocational training); or (3) higher secondary education (eight or nine years, leads to the *Abitur*, the university entrance degree). Sorting between these tracks is based on ability and teacher recommendations, but in some states parents may overrule teacher recommendations. It is a special aspect of the German educational system that individuals can switch between these tracks. Students who start off on a non-academic track have more than one chance to switch to higher secondary education in order to pursue, ultimately, university education.

Some students also switch from the academic track to the non-academic track. And it is not uncommon for students with a university entrance degree to opt for vocational training instead of a university education after earning their *Abitur*. Therefore, it seems appropriate to model educational choices as a choice about years of education rather than the schooling track. Education in Germany is generally tuition free at all stages.<sup>2</sup> Students from low economic background receive an allowance from the state (Bafög). The amount of that allowance depends on parental income and a share of about 50% must be repaid.

Table 1: SUMMARY OF EDUCATIONAL REFORMS

reform introduction	reform content
Phase 1: Academic track expansion	
1968	Implementation of advanced technical colleges
1972	Reorganization of the upper secondary level
1973	Implementation of integrated comprehensive schools parallel to the three-tier school system
Phase 2: Adapting to European schooling standards	
1999	Signing of Bologna Declaration marks transition from former 5-year university programmes to 3+2 years
2005	Implementation of Bologna Declaration structures
2007	Shortening of upper secondary level education by one year

The cohorts analyzed in our study were affected by two major educational reform phases, summarized in Table 1 that are crucial to our identification strategy. The first phase (1960s and 1970s) was devoted to the objective of equal opportunities for children from different socioeconomic backgrounds. It started with an expansion of universities in 1965 and the creation and introduction of universities of applied sciences in 1968. It continued with the introduction of Bafög in 1971 and the introduction of integrated comprehensive schools parallel to the standard three-tier school system in 1973.<sup>3</sup> Overall, this phase led to an expansion of the academic track. The second reform phase started in 2005 with the implementation of the Bologna process. Before this reform, the usual university academic degree was the *Diplom*, corresponding to about ten semesters of university education. The Bologna process introduced the Bachelor's degree as an intermediate academic degree and the Master's degree, which replaced the *Diplom*. The second phase continued with a decrease in the years of higher secondary education from nine to eight. Overall, this phase led to a reduction in schooling time.

<sup>2</sup>An exception is the period between 2005 and 2013 when universities in some states charged tuition fees of about 500 EUR per semester.

<sup>3</sup>Instead of opting for one of the tracks in the three-tier system, students of any ability can enroll in an integrated comprehensive school that combines all three options in one school without strictly separating students into education tracks.



### 3 Model

The basic model setup is drawn from Belzil and Hansen (2002). Individuals have rational expectations and maximize their present discounted value of expected lifetime utility by making annual schooling choices. Each year, they decide between continuing to go to school and exiting. Exit from school is defined as an absorbing state. Thus, this classifies as an optimal stopping problem. We distinguish between actual years of schooling and successful years of schooling. That is, when an individual invests an additional year at school, she faces uncertainty regarding the success of this year. By assumption, individuals leave school at the latest after 26 actual years in education or after having earned a Master's degree at university (highest degree observed), corresponding to 18 successful years of schooling.

At the beginning of each period  $t$ , the individual's information set consists of her age, cohort, successful years of schooling at time  $t$ , and region. There are two sources of uncertainty. On the one hand, the decision maker faces uncertainty over the success of one additional year of schooling. On the other hand, she faces uncertainty over her returns to schooling when she decides to exit education. The model accounts for both unobserved heterogeneity in schooling preferences and returns to schooling. We abstain from an explicit implementation of labor market processes. Instead, we exploit information on life cycle income profiles by successful years of schooling. Individuals form expectations over the distribution of these profiles according to their successful years. After exiting school, they receive a random draw from the respective distribution of lifetime income profiles.

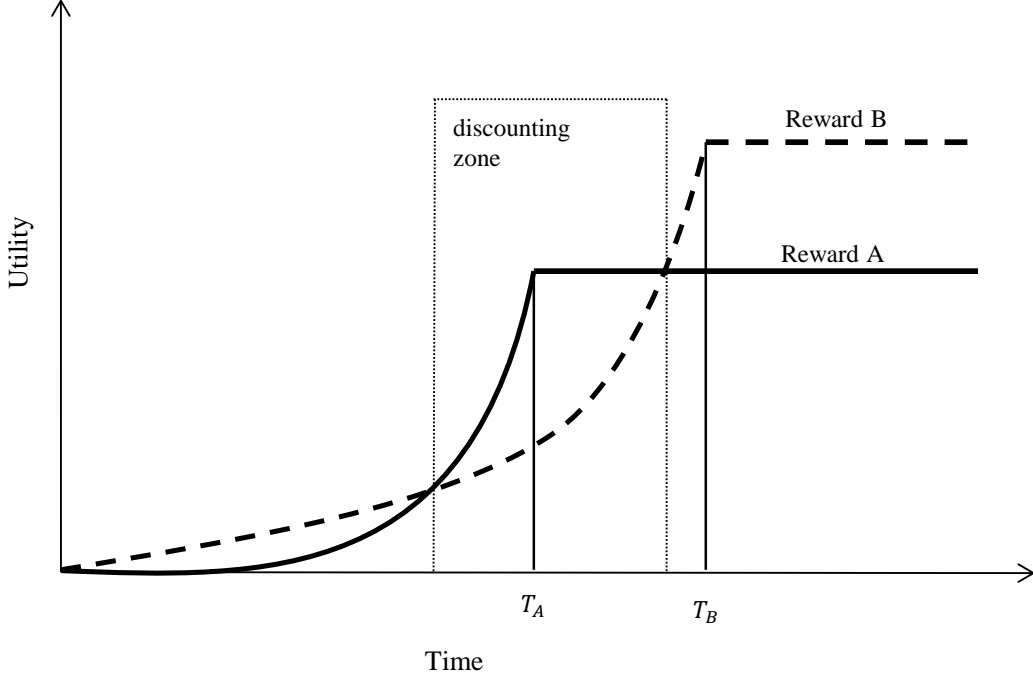
#### 3.1 Objective Function and Hyperbolic discounting

Individuals maximize their present discounted expected utility streams over the life-cycle. They are indexed by  $n$  and discrete time, the agent's age, is indexed by  $t$ . The decision period ranges from age 16 to 33, which corresponds to the interval of actual schooling years [9;26]. We model individual utility from age 16 to  $T = 70$ . The utility flow at age  $t$ ,  $U(S_{n,t}, d_{n,t})$ , depends on a vector of state variables,  $S_{n,t}$ , that affects the flow utilities and the agent's choice to continue school ( $d_{n,t} = 1$ ) or not ( $d_{n,t} = 0$ ). The important distinctive feature of this model is hyperbolic discounting that is captured by the following objective function:

$$EU_t = U(S_{n,t}, d_{n,t}) + \beta E_t \left[ \sum_{j=t+1}^T \delta^{j-t} U(S_{n,j}, d_{n,j}) \right] \quad (1)$$

We distinguish between a short-term discount factor  $\beta$ , also known as the *present-bias factor*, and a long-term discount factor  $\delta$  (O'Donoghue and Rabin, 1999). Be-

Figure 1: Illustration of Hyperbolic Discounting



solid: utility from reward A, dashed: utility from reward B

tween the current and the next period the individual discounts with  $\beta\delta$ , while she discounts with  $\delta$  between any two future adjacent periods. This is known in the literature as  $\beta$ - $\delta$ -preferences (Laibson, 1997). The standard model with exponential or time-consistent discounting is a special case of  $\beta$ - $\delta$ -preferences with  $\beta = 1$  while the individual exhibits hyperbolic discounting if  $\beta \in (0, 1)$ .

Figure 1 illustrates the way a hyperbolic discounter evaluates rewards over time. We consider two rewards in the future, a sooner low reward A and a later high reward B. At an early point in time, the time delay between both rewards does not affect the utility derived from both rewards such that  $B \succ A$ . However, at a critical proximity to the reward A the discounter reverses the preference relationship. Since preferences for the same rewards change over time, hyperbolic discounting is a form of time-inconsistent preferences.

The literature on hyperbolic time preferences distinguishes between *naïve* and *sophisticated* decision makers, who differ in the perception of discounting behavior for future periods (Strotz, 1955; Pollak, 1968; O'Donoghue and Rabin, 1999; O'Donoghue and Rabin, 1999). The sophisticated discounter is aware of the fact that she will also have present-biased preferences in future periods while the naïve discounter remains ignorant of this fact. Unlike the sophisticated hyperbolic discounter, the naïve

hyperbolic discounter is not aware of her future period self's self-control problem. Therefore, she has no intrinsic motivation to commit. In this case, only a binding commitment device imposed by a third party, e.g. the government, will provide incentives that make her behave in a way that is potentially more beneficial in the long term. In the context of educational policies, there is not much potential for commitment devices. The sophisticated discounter would simply discount future benefits more heavily than the naïve discounter and, thus, a comparison of the two types of hyperbolic discounters would not be very interesting. Within the scope of this study, we focus on naïve hyperbolic discounting.

### 3.2 Utility Function

Utility depends on consumption ( $C$ ) and preferences for schooling ( $\phi_{2g}$ ), where schooling preferences are heterogeneous and follow a Heckman and Singer (1984) type mass-point distribution that comprises three mass-points  $\kappa_g$  with  $g \in \{1, 2, 3\}$  and  $\kappa_1 + \kappa_2 + \kappa_3 = 1$ . We allow for correlation between schooling preferences and both the intercept and the returns to successful years of schooling in the income equation ( $\alpha_{0g}$  and  $\alpha_{1g}$ ). Note that schooling preferences and consumption are non-separable. That is, consumption affects utility through schooling-preferences as well as independently with a weight  $\phi_1$ . We also allow schooling-preferences to change separately from consumption by  $\phi_3$  after 13 years of schooling, the usual time period, after which individuals enter university. The utility function exhibits constant relative risk aversion (CRRA) and is separable across time:

$$U_{n,t} = \exp(\phi_1 + \phi_{2g}\mathbb{1}[d_{n,t} = 1]) \frac{C(S_{n,t}, d_{n,t})^\rho - 1}{\rho} + \phi_3 \mathbb{1}[sc_{n,t}^y \geq 14] + \epsilon_{n,t}(d_{n,t}) \quad (2)$$

where  $\epsilon(d_{n,t})$  follows a type 1 extreme value distribution. In our model we set  $\rho = -0.5$ , which yields a coefficient of relative risk-aversion of 1.5. This is consistent with previous evidence (see e.g. Blundell et al. (1994) and Chetty (2006)) and our results are insensitive to changes in the value of  $\rho$ <sup>4</sup>.

#### 3.2.1 Consumption after Schooling Exit

Instead of only considering earnings as returns to human capital, individual income is defined as a share of actual household income adjusted by the OECD modified equivalence scale.<sup>5</sup> This definition of income provides the advantage that individ-

<sup>4</sup>See Table 7 in section C for sensitivity checks of time preference estimates in relation to different values of  $\rho$

<sup>5</sup><http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf>.

uals' expectations regarding future income are not exclusively formed about labor market returns but also account for the sociological effect of educational choices. For instance, studies show that individuals are more likely to match with a partner with the same educational level (Mare, 1991). That is, even in the case of a highly educated person being unemployed, the higher education could still generate a higher income through the increased probability of being in a relationship with an equally well educated person who earns a higher income for both.

Individuals' beliefs on income after school are based on potential lifetime income profiles conditional on the number of successful years of schooling ( $sc_n^d$ ). In line with the heterogeneity in schooling preferences, the intercept of the income equation ( $\alpha_{0g}$ ) and the returns to successful years of schooling ( $\alpha_{1g}$ ) follow a mass-point distribution with three mass-points. After exiting school, individuals receive a draw from the distribution of lifetime income profiles that are derived from the following equation:

$$\log(inc_{n,t}) = \alpha_{0g} + \alpha_{1g}sc_n^d + \alpha_2age_{n,t} + \alpha_3age_{n,t}^2 + \eta_{n,t} \quad (3)$$

where  $\eta_{n,t}$  is a normally distributed random error.  $\alpha_1$  represents the returns to successful years of schooling and  $\alpha_2$  and  $\alpha_3$  describe the income-age profile. We set the existential minimum of  $inc_{n,t}$  to the average social security minimum of the year 2010 because all monetary values are price-adjusted to the year 2010.

### 3.2.2 Consumption at School

Similarly to consumption after schooling exit, consumption at school is derived from household income, adjusted according to the OECD modified equivalence scale. For children below 18 we adjust the adult share by 0.6 which corresponds to the OECD equivalence scale-adjustments of household income for children. The corresponding function is denoted by  $f(\cdot)$  such that consumption at school is given by

$$C(d_{n,t} = 1) = \begin{cases} f(HHinc_{n,t}) & \text{if } f(\cdot) \geq B \\ B & \text{if } f(\cdot) < B \end{cases} \quad (4)$$

The lower bound  $B$  is set according to the basic rules of the German state grant system for students (Bafög). We set  $B = 585$  which is the existential Bafög minimum in 2010. The idea of Bafög is that each student should have enough net income to maintain an existential level, where this level is defined by the national government to a specific monthly amount that has been increased infrequently since its implementation.

### 3.3 Transitions in Successful years of schooling

The number of actual years of schooling increases by one for every decision  $d_{n,t} = 1$ . Whether or not an additional year spent in education translates into a degree is subject to uncertainty. The probability to attain a degree and, thus, a higher value of successful years of schooling is derived from the latent variable  $sc_{n,t+1}^{d*}$ , which is given by the following equation:

$$sc_{n,t+1}^{d*} = \gamma_0(sc_{n,t}^y - sc_{n,t}^d) + \gamma_1 cohort2_n + \gamma_2 cohort3_n + \gamma_3 South_n + \gamma_4 South_n * cohort2_n + \gamma_5 South_n * cohort3_n + \zeta_{n,t} \quad (5)$$

where  $sc_{n,t}^d$  represents the number of successful years of schooling and  $sc_{n,t}^y$  represents the number of actual years of schooling.  $cohort2_n$  and  $cohort3_n$  are dummy variables indicating whether or not individual  $n$  is in the respective cohort group where the groups are defined as shown in Table 2. The cohort groups capture the reform phases of the German education system, as discussed in the previous section on the institutional framework. The regional indicator is represented by  $South_n$  and is equal to one if individual  $n$  lives in the south and 0 otherwise. We also include the respective interaction between  $South_n$  and the cohort dummy variables to account for the regional variation in the implementation of educational reforms.

We summarize the variables entering the transition equation (5) with  $S^\Pi$  as the set of state variables relevant for the transition probabilities. Given that  $sc_{n,t}^d$  has  $M = 11$  distinct values, we estimate  $K = 10$  corresponding cut-off points:  $m_1, \dots, m_K$ . Thus,

$$sc_{n,t+1}^d = \begin{cases} 1 & \text{if } sc_{n,t+1}^{d*} < m_1 \\ k & \text{if } m_{k-1} \leq sc_{n,t+1}^{d*} < m_k \\ 11 & \text{if } sc_{n,t+1}^{d*} > m_K \end{cases}$$

From this we can compute  $\Pi_m$  as a vector describing the discrete probability distribution for  $sc_{n,t+1}^d$  that results from equation (5). The parameters of equation (5)

Table 2: DISTRIBUTION OF COHORT GROUPS

cohort group	Perc %
1945-1963	41.38
1964-1984	50.64
1985-1992	7.97
<b>Total</b>	100.00

are summarized in the vector  $\gamma$  and we compute this probability in an ordered logit

model as follows:

$$Pr(sc_{n,t+1}^d > sc_j^d) = \frac{\exp(X\gamma - m_j)}{1 + \exp(X\gamma - m_j)}, j = 1, 2, \dots, M - 1, \quad (6)$$

such that

$$\begin{aligned} Pr(sc_{n,t+1}^d > sc_1^d) &= 1 - \frac{\exp(X\gamma - m_1)}{1 + \exp(X\gamma - m_1)} \\ Pr(sc_{n,t+1}^d = sc_j^d) &= \frac{\exp(X\gamma - m_{j-1})}{1 + \exp(X\gamma - m_{j-1})} \\ &\quad - \frac{\exp(X\gamma - m_j)}{1 + \exp(X\gamma - m_j)}, j = 2, \dots, M - 1 \\ Pr(sc_{n,t+1}^d = sc_M^d) &= \frac{\exp(X\gamma - m_{M-1})}{1 + \exp(X\gamma - m_{M-1})} \end{aligned} \quad (7)$$

When constructing  $\Pi$ , we need to account for the fact that the value of  $sc^d$  cannot decrease, such that  $sc_{n,t+1}^d \geq sc_{n,t}^d, \forall n, t$ .<sup>6</sup> Both the cohort and the regional variables are essential for the identification of time preferences. Their role as exclusion restrictions is discussed in section 4.

### 3.4 Value Functions: Hyperbolic Discounting

In this section, we discuss the construction of the expected value functions. For readability, we suppress the transition array in this section and use a uniform matrix of state variables,  $S$ . We assume that state transitions follow a Markov process that, by Bellman's principle of optimality, allows us to break down the value function to a two-period decision problem between the current and the present discounted value of future expected utility. In the basic exponential case the value function exhibits the form

$$V_t(S_{n,t}, d_{n,t}) = U(S_{n,t}, d_{n,t}) + \delta E_t[\max\{V_{t+1}^1, V_{t+1}^0\}] \quad (8)$$

where

$$\begin{aligned} V_{t+1}^1 &= V(S_{n,t+1}, d_{n,t+1} | S_{n,t} = s_{n,t}, d_{n,t} = 1) \\ V_{t+1}^0 &= V(S_{n,t+1}, d_{n,t+1} | S_{n,t} = s_{n,t}^-, d_{n,t} = 0) \end{aligned} \quad (9)$$

Following the terminology of Fang and Wang (2015), we call equation (8) the *perceived long-term value function*. Since the discount factor is constant over time in the

---

<sup>6</sup>We adjust the array accordingly, such that the remaining probabilities over the potential outcomes for  $sc_{n,t+1}^d$  still sum up to one. This is accomplished by taking the sum over all probabilities for values with  $sc_{n,t+1}^d < sc_{n,t+1}^d$  and adding equal shares of this sum to all remaining probabilities.

exponential discounting case, there is no difference between the decision maker's perception of her discounting behavior in the future and how she discounts in the present. Hence, expected payoffs over the short- and long-term are discounted equally. This leads to one value function covering all aspects of forward looking behavior.

For hyperbolic discounting, we distinguish between two aspects of forward looking behavior: 1) The individual's *actual* discounting behavior for the short- and long-term. 2) The individual's *perceived* discounting behavior for the short- and long-term. Therefore, we introduce a second value function that describes the discounting behavior of the hyperbolic discounter for the short-term:

$$W_t(S_{nt}, d_{nt}) = U(S_{nt}, d_{nt}) + \beta \delta E_t[\max\{V_{t+1}^S, V_{t+1}^W\}] \quad (10)$$

We call this function the *current value function*. If  $0 < \beta < 1$ , all non-immediate periods are discounted more heavily than in the exponential discounting case. Therefore, this type of discounting is also referred to as *present bias*. The case of exponential discounting is embedded in this model as a corner solution.

With hyperbolic discounting, the decision problem expands from a two-period to a three-period problem because we also consider how the current period self perceives the next period self to discount subsequent utility flows.

## 4 Identification of Time Preferences

Rust (1994) shows that the discount factor in dynamic discrete choice models with an infinite horizon is generally not identified. While further restrictions (e.g. absorbing state or finite horizon) help with the identification of parameters in the utility flows, Magnac and Thesmar (2002) show that these assumptions are not sufficient for the identification of time preferences. However, they also show that the non-identification issue can be resolved with certain exclusion restrictions. Furthermore, Fang and Wang (2015) derive the requirements for exclusion restrictions in order to identify additional time preference parameters that come with a model of hyperbolic discounting. Formally, the exclusion restriction has to meet the following conditions:

*There exist state variables  $x_1 \in S^\Pi$  and  $x_2 \in S^\Pi$  with  $x_1 \neq x_2$  such that:*

- (i) for all  $j \in J, U_j(x_1) = U_j(x_2)$
- (ii) for some  $j \in J, \Pi(S^{\Pi'}|x_1, j) \neq \Pi(S^{\Pi'}|x_2, j)$

where  $j \in J$  represents the choice and  $\Pi$  represents the transition probability matrix for the state variables. That is, there needs to be one variable that does not affect the current utility flows for any  $j \in J$  but affects the choices through its effect on

the transition probabilities of state variables. The rationale behind this identification strategy is as follows: if two individuals with the same current utility flow exhibit different value functions and, therefore, make different choices, this must be due to differences in their expectations about future outcomes.

The differences in the individuals' choices at different values of the exclusion restrictions provide information about the discount factor. For each additional parameter that describes discounting in the model, one additional exclusion restriction is required (Fang and Wang, 2015). Furthermore, identification relies on a setting with at least three consecutive decision periods. This is due to the fact that under hyperbolic discounting the value function in a dynamic choice model expands from a two- to a three-period problem. The behavioral difference between hyperbolic and exponential discounting is exhibited at the third period from the decision point. This is because for the exponential discounter the discount factor between period two and three is the same as between period one and two. However, for the hyperbolic discounter the discount factor is different at these points in time. This means that while we are able to identify the value of the discount factor in a two-period setting, we cannot distinguish hyperbolic discounting from exponential discounting with a low discount factor. Behavioral differences between these two types of discounting are only revealed through the differences in the discounting of two consecutive future periods. Using a panel data set with more than two periods and sufficient exclusion restrictions, we are able to identify this behavioral deviation.

The identification of our model relies on the assumption that the cohort group indicators (interacted with a regional indicator) affect the transition probabilities from actual to successful years of schooling, but are unrelated to the consumption and schooling preferences in the utility flows. The first assumption is not problematic because the cohort groups were subject to substantial changes in the educational policies. These changes affected both the average time spent on the attainment of specific degrees as well as the probabilities of completing schooling tracks successfully. The different policy regimes are described in section 2 and summarized in Table 1. In addition to the inter-cohort variation, we also exploit regional differences in the implementation of different reforms by including an indicator for southern states and its interaction with the cohort group dummies. The second assumption requires that the unobserved schooling preferences are individual-specific, but neither cohort- nor region-specific. Thus, we need to assume that the distribution of the schooling preferences does not differ across cohorts or region.



## 5 Solving the Model and Estimation

Given the finite horizon of the choice problem, the solution can be computed by backwards induction starting from the utility flow in the last decision period. After exit from schooling, the final value is defined through the distribution of lifetime income profiles. This yields the expected choice specific exit value of

$$E_t[W_{n,t}^0] = \beta \sum_{j=t+1}^T \delta^{j-(t+1)} U_{n,j}^0(S_{n,t} = \bar{s}_{n,t}, d_{n,t} = 0) \quad (11)$$

$E_t[W_{n,t}^0]$  represents the present discounted sum of future utility streams that evolve deterministically with age up to age  $T = 70$ . It follows from the type I extreme value distribution of  $\epsilon(d_{n,t})$  that we can derive a closed form solution for the expected maximum of future choice specific value functions (Rust, 1987):

$$E_t[W_{n,t}^1] = U(S_{n,t}, d_{n,t}) + \Pi(S_{n,t}^\Pi) \beta \delta \times \log\{\exp(E_t[V_{t+1}^1]) + \exp(E_t[V_{t+1}^0])\} \quad (12)$$

where,

$$E_t[V_{n,t+1}^1] = U(S_{n,t+1}, d_{n,t+1}) + \Pi(S_{n,t+1}^\Pi) \delta \times \log\{\exp(E_{t+1}[V_{t+2}^1]) + \exp(E_{t+1}[V_{t+2}^0])\} \quad (13)$$

For the backward induction with naïve hyperbolic discounting, we compute the value functions recursively until  $t + 1$  using the perceived long-term value function and compute the last step from  $t + 1$  to  $t$  using the current value function.

Rust (1987) shows that with the assumption of additive separability in utility over time and conditional independence, the probability of a continuation of schooling in period  $t$  takes on the following logit-type form:

$$Pr(d_{n,t} = 1) = \frac{\exp(E[W_{n,t}^1])}{\exp(E[W_{n,t}^0]) + \exp(E[W_{n,t}^1])} \quad (14)$$

The probability to remain in school for  $\bar{t}$  years (after grade 9) for individual  $n$  is:

$$L_{1n} = \sum_g \kappa_g \left\{ \prod_{t=1}^{\bar{t}} Pr(d_{n,t} = 1) \times Pr(d_{n,\bar{t}+1} = 0) \right\} \quad (15)$$

The full log-likelihood of the model can be obtained by multiplying the likelihood contribution from (15) with the income density that is computed from the income

equation (3) ( $L_{2n}$ ).

$$\sum_{n=1}^N \{ \log(\sum_g \kappa_g \{L_{1n} \times L_{2n}\}) \} \quad (16)$$

We compute the transition probabilities presented in equation (5) ( $L_{3n}$ ) in a first step and take them as given in the main estimation procedure. The parameters of our model are estimated by the method of maximum likelihood.

## 6 Data

This study is based on the German Socio-Economic Panel (SOEP), an annual survey that, since 1984, collects individual- and household-level information from about 12,000 households (Wagner et al., 2007). We confine our analysis to the 1992-2013 waves (collected after the German reunification) and focus on West German students with a West German school degree. This allows us to abstract from institutional differences in former East German regions. Furthermore, we exclude disabled individuals and individuals who remained in education for more than 26 years (outliers that account for less than 2% of the sample). Since missing income information must be imputed, we further reduce the sample to individuals with a minimum of two observed periods of household income during the educational period. The final sample is a balanced panel consisting of 2305 individuals.

Table 3: Summary Statistics

	Mean	Std.	Min	Max
age	24.91	4.56	16.00	33.00
successful years of school ( $sc^d$ )	13.64	2.58	9.00	18.00
actual years at school ( $sc^y$ )	15.76	3.35	9.00	26.00
equivalent income p.a.(10,000 €)	1.84	1.03	0.79	22.68
net HH income p.a.(10,000 €)	3.33	2.36	0.79	56.71
cohort2	0.37	0.48	0.00	1.00
cohort3	0.46	0.50	0.00	1.00
south	0.57	0.49	0.00	1.00

Note: Authors' calculations based on SOEP.

Table 3 shows summary statistics of the relevant variables. Note that we distinguish between actual years of schooling and successful years of schooling. We use the following variables in the estimation of our model:

**Actual years of schooling** Actual years of schooling are defined as the last age that the individual is observed in any educational institution minus seven (the age most students reach during their first year at school). The variable ranges from 9 to 26 years (9 years is the minimum number of years required by compulsory schooling laws and 26 years is the highest number of school years that we observe, covering more than 98% of our sample). We drop a 2%-share of individuals who need more than six years longer at school than the regular years for a certain degree.

**Successful years of schooling** Successful years of schooling are derived from information on obtained degrees. Although we do not observe the assigned schooling variable in each period, the supplemental biographical data set *BIOEDU* (Lohmann and Witzke, 2011) provides information on educational participation and transitions of the individuals in the SOEP. We use this to reconstruct educational paths. We then assign years of schooling to the respective degrees. The method used to assign years of schooling to a degree follows Couch (1994).

**Income** Individual income is computed by adjusting total net income (also covering government transfers) of all household members according to the OECD modified equivalence scale.<sup>7</sup> We set the minimum annual income of an individual to the average social security minimum in 2010<sup>8</sup> in order to avoid any bias due to unobserved transfers in the income variable. Figure 7 in the appendix shows the average annual income by different levels of education. Similarly to labor market earnings profiles, these income profiles are hump-shaped and differ by levels by education.

**Cohort groups** The cohorts range from 1946 to 1992. We define three cohort groups as follows: group 1 consists of all born before 1964, group 2 consists of all born between 1964 and 1984, while group 3 consists of everyone born after 1984. It is not easy to specify cohort groups that were clearly affected by one or the other educational reform phase. This is due to the differences in years of schooling invested for specific degrees but also due to the fact that the duration of both reform phases spans several years. Nevertheless, we chose this categorization for the following reasons: the oldest cohorts of group 2 were nine years old and, thus, mostly in the fourth grade when the implementation of the integrated comprehensive schools started. This is the age when the decision for the secondary schooling track is made. The oldest cohorts of group 3 were 20 years old when the implementation of the Bologna process started. This is the age when most students enter university.

**South** This variable is equal to 1 if the individual lives in a Southern state and 0 otherwise. The southern states are Hesse, Rhineland-palatinate, Baden-Wuerttemberg, Bavaria and Saarland. The remaining states are Schleswig-Holstein, Hamburg, Lower-Saxony, Bremen and North-Rhine-Westphalia.

---

<sup>7</sup><http://www.oecd.org/eo/growth/OECD-Note-EquivalenceScales.pdf>.

<sup>8</sup>All monetary values in this study are price adjusted to 2010.

## 7 Estimation Results

In the following, we estimate two versions of our model: one with hyperbolic and one with exponential discounting. Under hyperbolic discounting, there are two parameters to explain observed discounting behavior. And under exponential discounting, there is one parameter to explain the same economic feature. When  $\beta < 1$ , we expect  $\delta$  to be lower in the exponential version than the hyperbolic version of the model because the model is used to rationalize the same degree of myopia. Note that the exponential model is nested within the hyperbolic model. Therefore, we can compare both models in their ability to explain the data in the framework of a likelihood-ratio test.

### 7.1 Transition Probabilities

Table 4: Ordered Logit Results of Transition Probabilities

	Value	Std.Err	T-Value	P-Value
$l.sc^y - l.sc^d$	0.5051***	0.0077	65.7346	0.0000
<i>cohort2</i>	-0.0370	0.7074	-0.0522	0.3984
<i>cohort3</i>	-0.2195***	0.0217	-10.1218	0.0000
<i>south</i>	0.0514***	0.0086	5.9731	0.0000
<i>south * cohort2</i>	-0.0543	0.7074	-0.0767	0.3978
<i>south * cohort3</i>	0.2338***	0.0136	17.2390	0.0000
No. of Cutpoints	10			
<i>N</i>	2305			
<i>ll</i>	29433.4135			

\*, \*\* and \*\*\* denote significance level of 10%, 5% and 1%, respectively.

The estimation results of the transition equation are reported in Table 4. A likelihood-ratio test against the hypothesis of zero impact of the exclusion restrictions yields a p-value of 0. However, the impact of the first reform phase is not significant and does not differ between northern and southern states. The negative coefficient for the *cohort3* dummy suggests that the last reform phase generally decreased the probability of attaining higher degrees for both northern and southern states whereas much less for the southern states. This supports the hypothesis that the regional variations stem from more resilient implementation of educational reforms in the south. Overall, these results support the relevance of the exclusion restrictions.

Table 5: Schooling and Utility Parameters

	Hyperbolic		Exponential	
	Value	Std.Err	Value	Std.Err
$\beta(shortrun)$	0.5696***	0.0147	-	-
$\delta(longrun)$	0.9743***	0.0491	0.9136***	0.0094
$\rho$	-0.5	-	-0.5	-
$\phi_1(C - weight)$	0.4587***	0.0466	0.6078***	0.0629
$\phi_3(sc^y \geq 14)$	-0.7162***	0.0351	-0.7919***	0.0322
$prob(type1)$	0.2212	-	0.5083	-
$prob(type2)$	0.5075	-	0.3501	-
$prob(type3)$	0.2712	-	0.1416	-
$\phi_{21}(type1 : schooldisut)$	-4.1969***	1.1673	-1.7903***	0.1634
$\phi_{22}(type2 : schooldisut)$	-1.4657***	0.1437	-3.9764***	0.6534
$\phi_{23}(type3 : schooldisut)$	-4.4377***	1.2880	-3.9287***	1.4772
$N$	2305	-	2305	-
$ll$	-43461.7359	-	-43485.9459	-
<b>LR-Test Exp vs. Hyp</b>	0.0000	-	-	-

\*, \*\* and \*\*\* denote significance level of 10%, 5% and 1%, respectively.

## 7.2 Flow Utility and Income

The estimation results for the parameters of the utility function are reported in Table 5. As mentioned before, we set  $\rho = -0.5$ , which corresponds to a coefficient of relative risk aversion of 1.5.

For the hyperbolic model, we estimate a short-term discount factor of 0.57 and a long-term discount factor of 0.97. For the exponential model, we estimate a discount factor of 0.91. We find strong heterogeneity in the disutility of schooling. In the hyperbolic model, type 3 exhibits a disutility that is about triple the size of the disutility of type 2. We only find moderate differences in the estimated parameters of the utility function between the exponential and the hyperbolic models.

Since the exponential model is nested within the hyperbolic model, we can test the exponential model against the hyperbolic model using a likelihood-ratio test. With a p-value of 0, we reject the null of the exponential model in favor of the hyperbolic model. Hence, this is strong empirical evidence of hyperbolic discounting in educational decisions.

Finally, the results for the income equation are presented in Table 6. We report returns to successful years of schooling of that range between 6.7% and 8.5% for the three types. These effects are at the lower bound of what is usually found in the

Table 6: Income Equation

	Hyperbolic		Exponential	
	Value	Std.Err	Value	Std.Err
$\sigma_\eta$	0.2906***	0.0014	0.2905***	0.0006
$\alpha_{01}(\text{type1} : \text{const.})$	6.7632***	0.0193	7.1773***	0.0183
$\alpha_{11}(\text{type1} : \text{sc}^d)$	0.0670***	0.0010	0.0735***	0.0008
$\alpha_{02}(\text{type2} : \text{const.})$	7.1736***	0.0179	6.7691***	0.0196
$\alpha_{12}(\text{type2} : \text{sc}^d)$	0.0733***	0.0008	0.0670***	0.0010
$\alpha_{03}(\text{type3} : \text{const.})$	7.5215***	0.0222	7.5188***	0.0196
$\alpha_{13}(\text{type3} : \text{sc}^d)$	0.0848***	0.0013	0.0854***	0.0013
$\alpha_3(\text{age})$	0.0661***	0.0034	0.0659***	0.0009
$\alpha_4(\text{age}^2)$	-0.0005***	0.0000	-0.0005***	0.0000

\*, \*\* and \*\*\* denote significance level of 10%, 5% and 1%, respectively.

empirical literature (Card, 1999), but it compares favorably to the effects found in the structural model of Belzil and Hansen (2002). This may be due to the choice of using net household-equivalent income instead of gross labor earnings as our income measure. Hence, the returns to education are lower because of the redistributive effects of the tax and transfer system.

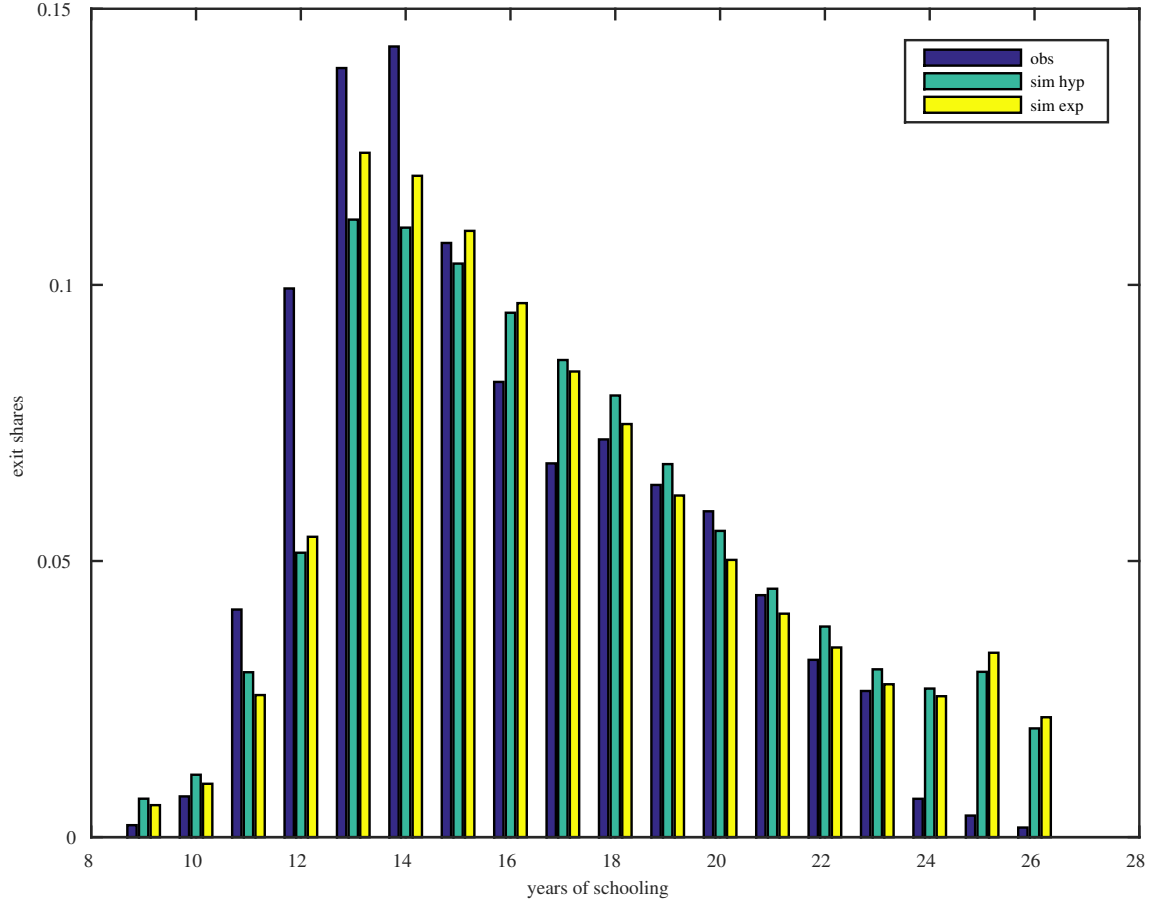
Our estimates of the coefficients on  $\text{age}$  and  $\text{age}^2$  indicate a concave relationship between  $\log(\text{inc})$  and  $\text{age}$ , as well established in the literature. Overall, our estimates of the income equation show reasonable results that fit well into the model and the existing literature.

### 7.3 Model Fit

We check the internal validity of not only our preferred model specification with hyperbolic discounting, but also the specification with exponential discounting. For this purpose, we compare simulated exit shares with the respective shares from the estimation sample. Figure 2 depicts the comparison. We simulate synthetic samples with  $N = 50,000$  individuals based on the point estimates of the parameters of the respective model specifications. For the cohort groups and the south indicator, we rely on the distribution observed in the estimation sample. Choices and transitions are simulated taking draws from the uniform distribution.

The model fits the overall shape of the observed distribution reasonably well. We somewhat underpredict the exit shares at the lower end of the distribution and overpredict them at the upper end. However, we refrain from improving the in-sample fit artificially by including more years of schooling dummies because it is ultimately our goal to rationalize choices based on a model of economic reasoning.

Figure 2: Educational Decisions: Observed and Simulated



Choice probabilities are non-parametrically estimated based on simulated samples.

The focus of this study lies on the comparison of behavioral responses to policy changes. Therefore, small deviations in the levels between the predicted and observed exit shares are very unlikely to affect these results.

## 8 Policy Analysis

We consider policies that affect individual utility streams at different points in time. In particular, we investigate the effects of an increase in the state-supplied student loans (BaföG). Generally, student loans are an internationally widespread policy measure to encourage take-up of higher education. They affect the intertemporal substitution of consumption and, therefore, represent an excellent setting for the analysis of time discounting. The regulatory structure of payback laws differ between countries. We simulate two different variants of this policy: (1) an increase in the state grant for students as a way to affect short-term costs while at school; and (2) an increase in the state grant as a loan that will have to be paid back after the end of the education. The increase of the student loan is implemented through an increase of the lower bound  $B$  in equation (4). Since income during the schooling period is a function of household income, not everyone is affected by this policy. Only individuals from lower income groups with an initial income below the threshold receive this loan.

In our analysis, we look at both the short-term effects in terms of changes in the educational choices and at the long-term effects in terms of changes in the life-cycle income. The policy measures affect educational choices through the relative utility costs of schooling and through an increase in the financial means available to the individuals. If the increase in the state grant is provided as a loan, the payback scheme affects the relative utility costs of schooling differently depending on how individuals discount the future repayments. The income returns to education are expected to differentially affect individuals under hyperbolic or exponential discounting.

### 8.1 Student Grant

In this policy variant, the short-term gain does not correspond to long-term costs. We simulate the increase in student loans by gradually raising the minimum threshold of income at school ( $B$ ) by 20€ per month.

We run all simulations for a synthetic sample of 50,000 simulated individuals. Figure 3 shows the effect of a 20 to 200 € increase in the monthly grant on actual years of schooling. The left part presents the average increase in years spent at school as a response to a higher monthly grant of 20 to 200€. The right part presents the share of individuals who stay longer at school as a response to the policy changes. The actual years spent at school increase under both model specifications. These results show that under hyperbolic discounting individuals react slightly stronger to the policy and stay longer at school. With respect to effect sizes, it can be said that a 100€ increase in the monthly grant leads to an average increase of about 0.03 years (or about 1/3 of a month) among the individuals who go to university. A share of



Figure 3: Effect of a higher grant on actual years of schooling

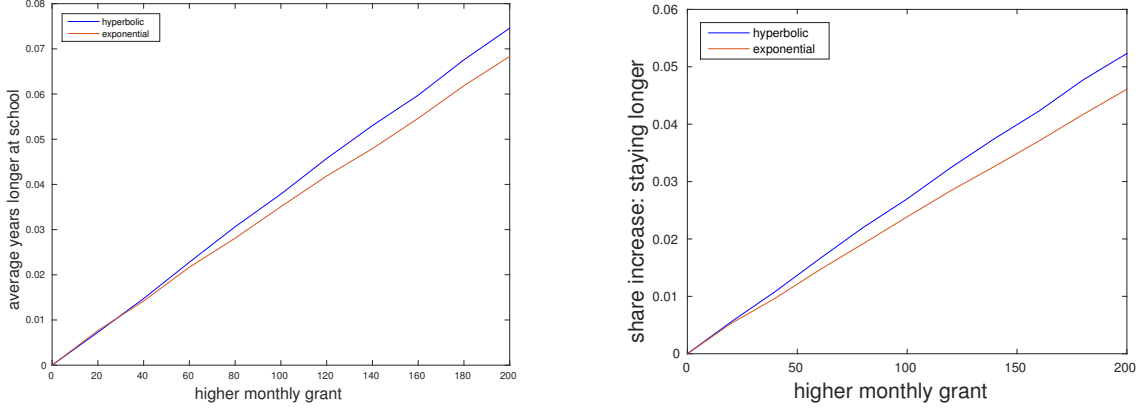
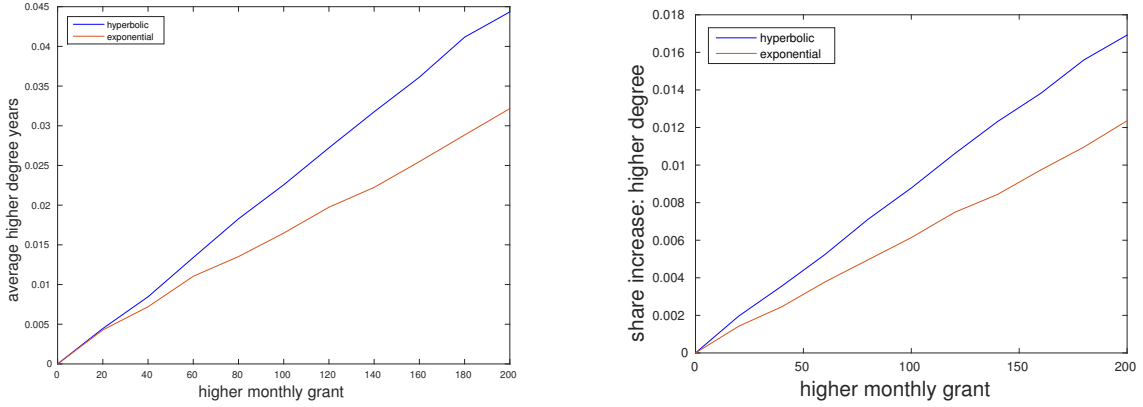


Figure 4: Effect of a higher grant on successful years of schooling



about 3% of the population stays longer in education. In terms of magnitude these estimates are in line with the effect sizes found by Steiner and Wrohlich (2012) using a discrete-time hazard rate model for Germany.

We observe stronger differences between exponential and hyperbolic discounting when we consider the effects of the policy on successful years of schooling as it is displayed in Figure 4. The effects are larger under hyperbolic discounting. This reflects the fact that the hyperbolic discounter behaves less efficiently and can realize higher gains by staying one year longer at school. For example, an individual is more likely to be at the margin of obtaining a degree when dropping out before the reform under hyperbolic discounting. Within the framework of our model, only additional successful years of schooling yield a higher income after school exit while actual years of schooling only translate into a higher probability of achieving additional successful years. The exponential discounter puts more weight on the specific transition probabilities of actual to successful years of schooling when making the choice (e.g. tends not to drop out shortly before obtaining a degree before the reform). This is in line with the behavioral specifications of the two types of discounting.

Figure 8 in the Appendix shows the effect of a 100€ increase in the monthly grant

on the exit shares by actual years of schooling. Under exponential discounting, the changes in actual years of schooling do not strongly affect the overall distribution of actual years of schooling with small changes at the tails of the distribution and larger changes in the middle of the distribution. Overall, this leads to a more equal distribution of actual years of schooling. Under hyperbolic discounting, individuals generally exhibit a similar behavior, however with a particularly strong decrease in the exit share for individuals with eleven years of schooling that corresponds to a large spike at 16 years of schooling. The probability to obtain a degree after eleven years of schooling is relatively small while it is comparatively larger after twelve or 16 years of schooling. This further explains the stronger effect on successful years of schooling under hyperbolic discounting.

## 8.2 Student Loan

Unlike the grant, the student loan needs to be repaid once the individual exits school. Individuals start repaying their loan after school exit in equal annuities over five years. Again, we run all simulations for a synthetic sample of 50,000 simulated individuals and gradually increase the provided loan from 20 to 200€. In this policy variant, the short-term gains correspond to long-term losses in consumption when the repayments start.

Figure 5 shows the effect of a 20 to 200€ increase in the monthly loan on actual years spent at school. For the purpose of comparison, we also add the results from the grant simulation to the plot. These results show that the response to the two policies differs more for exponential than for hyperbolic discounters.

Figure 6 shows the effect of the policy on successful years of schooling. Again, we also plot the effects of the grant policy in dashed lines. Similar to the grant policy, we see larger effects on the successful years of schooling for the hyperbolic discounter. This can also be explained by the more efficient behavior of the exponential discounter as opposed to the more impulsive behavior of the hyperbolic discounter before the reform. In both cases, the response to the loan policy is stronger than for the grant. It appears that individuals try to offset the negative effects on consumption in the long-term by earning higher degrees. The payback mechanism triggers a further positive effect on school participation. This is, however, less the case for the hyperbolic discounters.

Figure 9 in Appendix B presents the effect of a 100€ increase in the monthly loan on the shares of students exiting school by actual years of schooling. Again, under hyperbolic discounting we observe a strong decrease at eleven years of schooling that corresponds to a large spike at 16 years. This also explains the strong effects on successful years of schooling under hyperbolic discounting. Under exponential

Figure 5: Effect of a higher loan on actual years of schooling

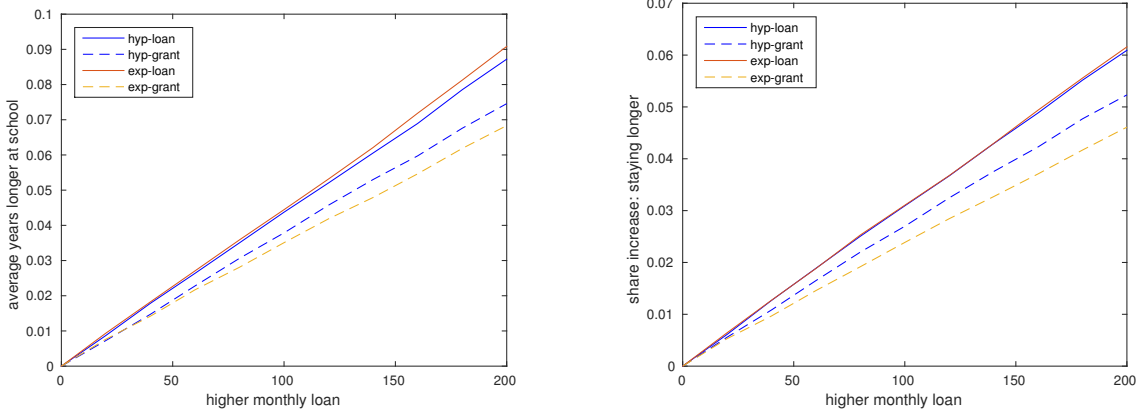
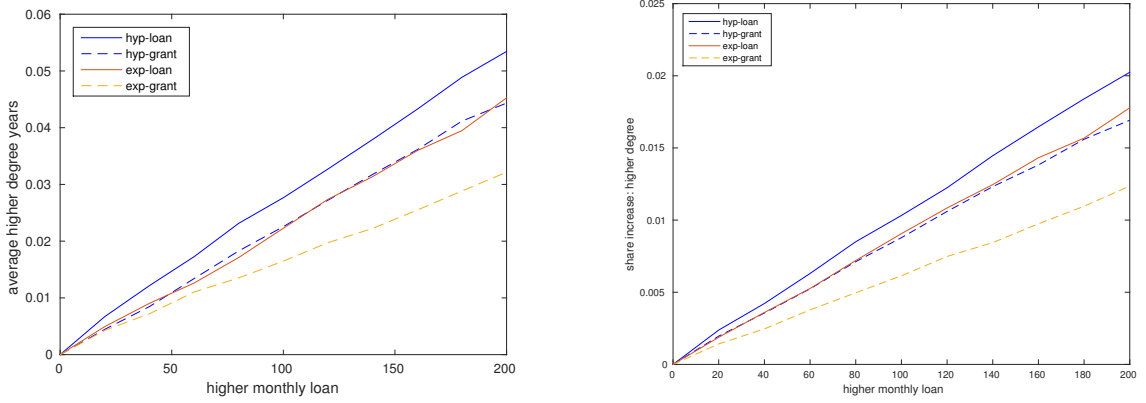


Figure 6: Effect of a higher loan on successful years of schooling



discounting, we also observe a large drop at twelve actual years (usually the time spent to achieve the university access degree) and somewhat larger spikes at 16 and 18 years (corresponding to the usual amount of years spent to achieve a bachelor's or a masters degree). This again supports the hypothesis that hyperbolic discounters are more often at the margin of obtaining a degree when leaving education, while exponential discounters more often just obtained a degree when they leave education.

### 8.3 Sensitivity Analysis

In our model we set  $\rho = -0.5$ , which yields a coefficient of relative risk aversion (CRRA) of 1.5. We check the sensitivity of this assumption by estimating the same model with different values of  $\rho$  that correspond to a CRRA parameter in the range  $[1, 2]$  (see Chetty (2006) for a discussion of the possible range of this parameter). The results are summarized in Table 7. In the hyperbolic model, the short-term discount factor  $\beta$  decreases (from 0.677 to 0.51) with rising degrees of risk aversion, while the long-term discount factor  $\delta$  increases (from 0.936 to 1.011). Similarly, in the exponential model  $\delta$  increases (from 0.886 to 0.941) with rising degrees of risk

aversion. That is, we observe an increase in the long-term discount factor as risk aversion rises. Despite the interdependence of these three parameters, the overall pattern and, therefore, our main results are very robust to the whole range of levels of relative risk aversion. In all cases, we see that the short-term discount factor is substantially smaller than the long-term discount factor in the hyperbolic model and that the long-term discount factor in the exponential model is smaller than in the hyperbolic model (offsetting the restriction of  $\beta = 1$  in the exponential model).

## 9 Conclusion

We investigate time-inconsistent preferences in educational decision making and corresponding policies using a structural dynamic choice model. Based on a novel identification approach, we exploit the variation in average years invested in degree attainment through various educational reforms to identify the discount factor of hyperbolic time preferences. We achieve identification by imposing exclusion restrictions that affect educational choices indirectly through their impact on the transition probabilities of relevant state variables but have no impact on current utility flows. The estimates indicate time-inconsistent behavior and provide quantitative evidence for its relevance. Failure to account for this behavioral feature in a structural model may lead to misleading results in subsequent policy simulations.

In our study, this is demonstrated for two different policy scenarios: (1) an increase in the state grant for students as a way to affect short-term costs while at school; and (2) an increase in the state grant as a loan that will have to be paid back after completing education. We find substantial differences in the effects of these policies when comparing educational outcomes based on a model specification with hyperbolic discounting with the ones based on a specification with exponential discounting. In particular, the effects on successful years of schooling are much stronger for hyperbolic discounters. This reflects the fact that the hyperbolic discounter behaves less efficiently and can realize higher gains by staying one year longer at school. Moreover, the response to the two policies differs more for exponential than for hyperbolic discounters.

## References

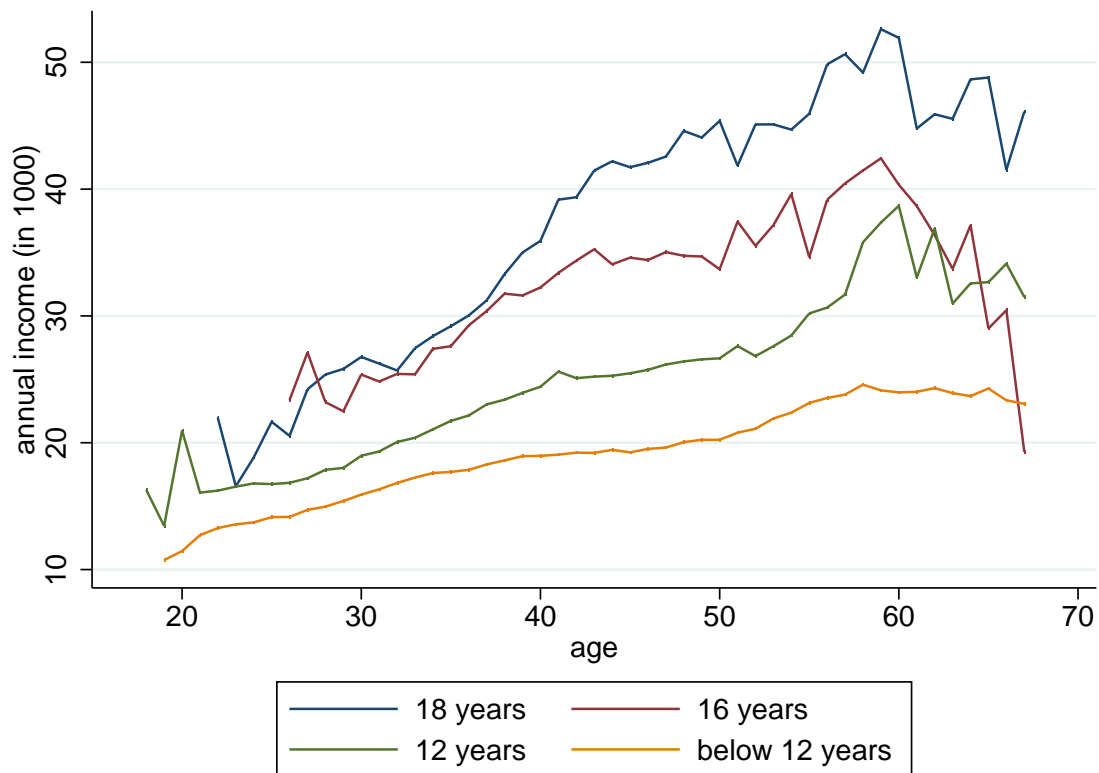
- Barro, R. J. (1998). *Determinants of Economic Growth: A Cross-Country Empirical Study*, Volume 1 of *MIT Press Books*. The MIT Press.
- Belzil, C. and J. Hansen (1999). Subjective discount rates, intergenerational transfers and the return to schooling. IZA Discussion Papers 60, Institute for the Study of Labor (IZA).
- Belzil, C. and J. Hansen (2002). Unobserved ability and the return to schooling. *Econometrica* 70(5), 2075–2091.
- Belzil, C. and M. Leonardi (2007). Can risk aversion explain schooling attainments? evidence from Italy. *Labour Economics* 14(6), 957 – 970.
- Blundell, R., M. Browning, and C. Meghir (1994). Consumer Demand and the Life-Cycle Allocation of Household Expenditures. *Review of Economic Studies* 61(1), 57–80.
- Cameron, S. and J. Heckman (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy* 109(3), 455–499.
- Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 30, pp. 1801–1863. Elsevier.
- Carneiro, P., K. T. Hansen, and J. J. Heckman (2003). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college. *International Economic Review* 44, 361–422.
- Chan, M. K. (2013). A dynamic model of welfare reform. *Econometrica* 81(3), 941–1001.
- Chetty, R. (2006). A new method of estimating risk aversion. *The American Economic Review* 96(5), 1821–1834.
- Cohen, J. D., K. M. Ericson, D. Laibson, and J. M. White (2016, July). Measuring time preferences. Working Paper 22455, National Bureau of Economic Research.
- Couch, K. A. (1994). High school vocational education, apprenticeship, and earnings: a comparison of Germany and the United States.
- DellaVigna, S. and M. D. Paserman (2005). Job search and impatience. *Journal of Labor Economics* 23(3), pp. 527–588.
- Eckstein, Z. and K. I. Wolpin (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica* 67(6), pp. 1295–1339.
- Fang, H. and D. Silverman (2006). Distinguishing between cognitive biases: Beliefs vs. time discounting in welfare program participation. *Behavioral Public Finance*, 47–81.

- Fang, H. and D. Silverman (2009, November). Time-inconsistency and welfare program participation: Evidence from the nlsy. *International Economic Review* 50(4), 1043–1077.
- Fang, H. and Y. Wang (2015). Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. *International Economic Review* 56(2), 565–596.
- Giné, X., D. Karlan, and J. Zinman (2010). Put your money where your butt is: a commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, 213–235.
- Gustman, A. L. and T. L. Steinmeier (2012). Policy effects in hyperbolic vs. exponential models of consumption and retirement. *Journal of Public Economics* 96(5–6), 465 – 473.
- Haan, P., L. Haywood, and U. Schneider (2016). Time preferences and female labor supply. *mimeo*.
- Heckman, J. and B. Singer (1984, March). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica* 52(2), 271–320.
- Heckman, J. J., J. E. Humphries, and G. Veramendi (2016, May). Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking. IZA Discussion Papers 9957, Institute for the Study of Labor (IZA).
- Heckman, J. J., L. Lochner, and P. E. Todd (2005, August). Earnings functions, rates of return and treatment effects: The mincer equation and beyond. IZA Discussion Papers 1700, Institute for the Study of Labor (IZA).
- Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy* 105(3), pp. 473–522.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 443–477.
- Laibson, D., A. Repetto, and J. Tobacman (2007, August). Estimating discount functions with consumption choices over the lifecycle. NBER Working Papers 13314, National Bureau of Economic Research, Inc.
- Lindahl, M. and A. B. Krueger (2001, December). Education for Growth: Why and for Whom? *Journal of Economic Literature* 39(4), 1101–1136.
- Lohmann, H. and S. Witzke (2011). Bioedu (beta version): Biographical data on educational participation and transitions in the german socio-economic panel study (soep). Technical report, DIW Berlin.
- Magnac, T. and D. Thesmar (2002, March). Identifying dynamic discrete decision processes. *Econometrica* 70(2), 801–816.
- Mare, R. D. (1991). Five decades of educational assortative mating. *American sociological review*, 15–32.

- Mischel, W., Y. Shoda, and M. L. Rodriguez (1989). Delay of gratification in children. *Science* 244(4907), 933–938.
- O’Donoghue, T. and M. Rabin (1999). Addiction and self-control. *Addiction: Entries and exits*, 169–206.
- O’Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review* 89(1), 103–124.
- Oosterbeek, H. and H. van Ophem (2000). Schooling choices: Preferences, discount rates, and rates of return. *Empirical Economics* 25(1), 15–34.
- Paserman, M. D. (2008). Job search and hyperbolic discounting: Structural estimation and policy evaluation. *The Economic Journal* 118(531), 1418–1452.
- Pollak, R. A. (1968). Consistent planning. *The Review of Economic Studies*, 201–208.
- Rust, J. (1987, September). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica* 55(5), 999–1033.
- Rust, J. (1994). Structural estimation of markov decision processes. *Handbook of econometrics* 4(4).
- Steiner, V. and K. Wrohlich (2012, 03). Financial Student Aid and Enrollment in Higher Education: New Evidence from Germany. *Scandinavian Journal of Economics* 114(1), 124–147.
- Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 165–180.
- Tobacman, J. and P. Skiba (2005, November). Payday Loans, Consumption Shocks, and Discounting. *Computing in Economics and Finance* 2005 189, Society for Computational Economics.
- Wagner, G. G., J. R. Frick, and J. Schupp (2007). The german socio-economic panel study (soep)-evolution, scope and enhancements.

## A Supplemental Descriptive Statistics

Figure 7: Annual average of income by different levels of education





## B Further Policy Analysis

Figure 8: Difference in exit shares by actual years of schooling for 100€ increase in a monthly grant

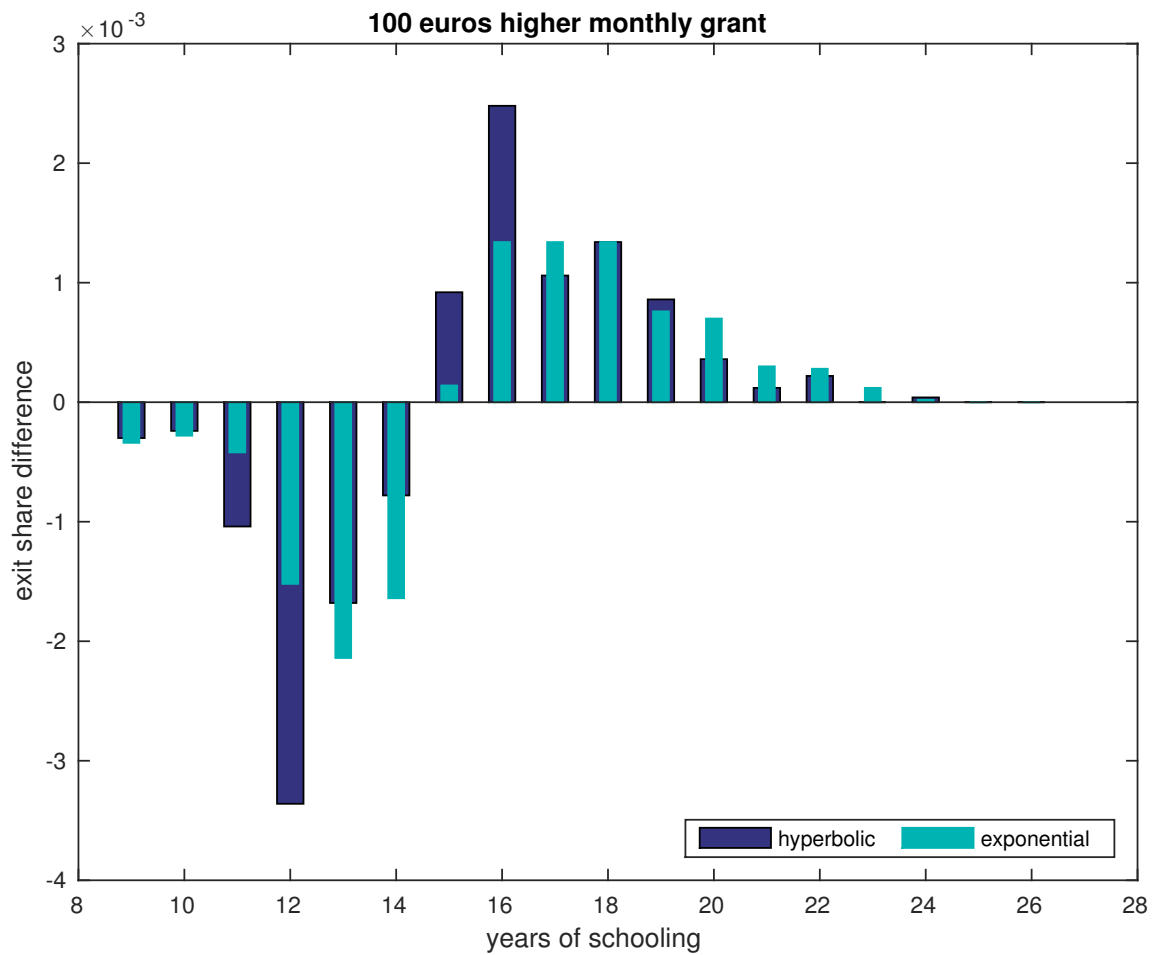
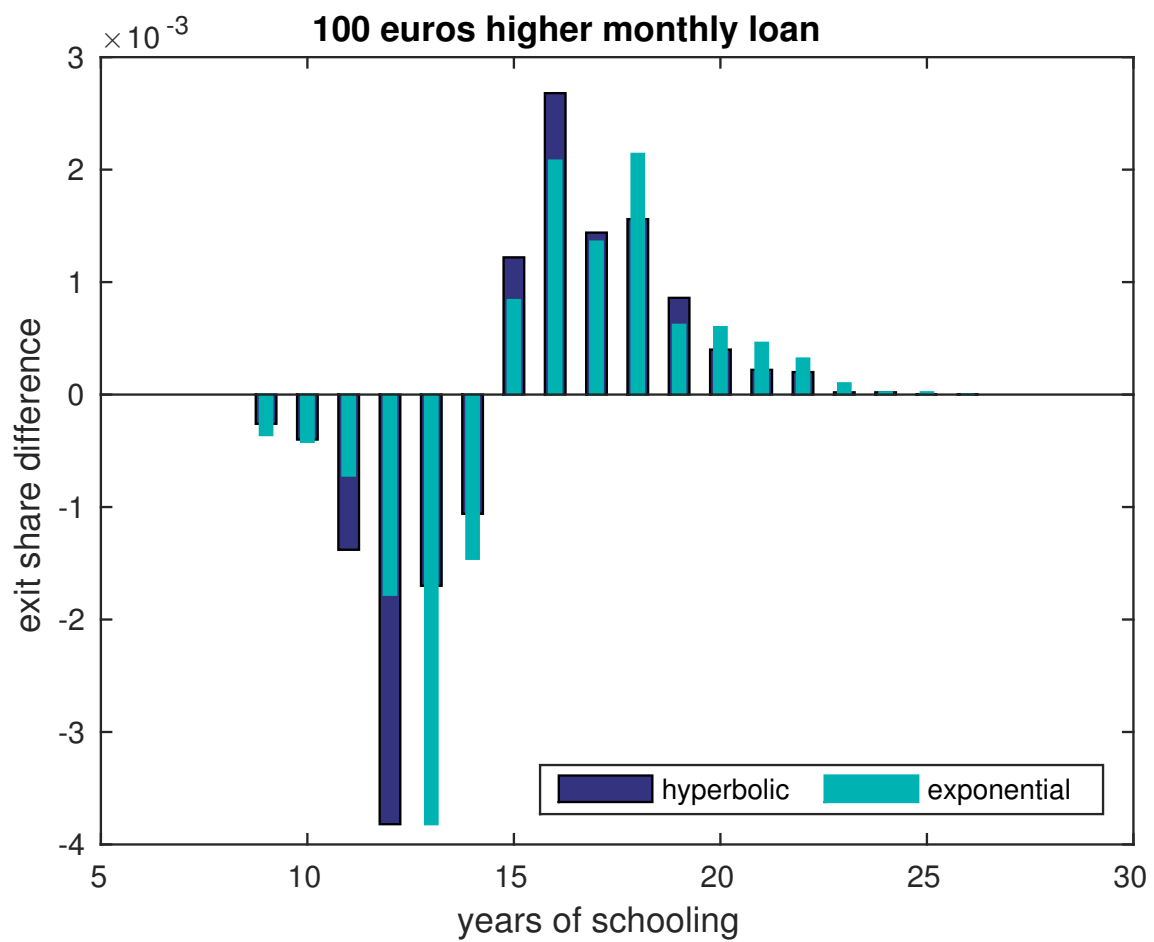


Figure 9: Difference in exit shares by actual years of schooling for 100€ increase in a monthly loan



## C Sensitivity Check

Table 7: Sensitivity of  $\beta$  and  $\delta$  with respect to  $\rho$

$1 - \rho$	Hyperbolic		Exponential
	$\beta$	$\delta$	$\delta$
1.0	0.677	0.936	0.886
1.1	0.677	0.936	0.898
1.2	0.603	0.959	0.898
1.3	0.603	0.959	0.908
1.4	0.540	0.984	0.908
1.5	0.570	0.974	0.914
1.6	0.540	0.984	0.925
1.7	0.519	0.999	0.925
1.8	0.519	0.999	0.936
1.9	0.510	1.011	0.936
2.0	0.510	1.011	0.941

Source: Own calculations. Estimation results of time preference parameters  $\beta$  and  $\delta$  for different values of the coefficient of relative risk aversion ( $1 - \rho$ ).