Intergenerational Mobility:
New Evidence from Consumption Data

Gustaf Bruze*

Karolinska Institute

May 3, 2015

Abstract

This paper constructs the permanent consumption of Danish parents and children from fifteen years of panel data and estimates the persistence of consumption across generations. The intergenerational elasticity of consumption is more than twice as high as the widely-reported intergenerational elasticity of earnings, and 35% higher than the intergenerational elasticity of disposable income. The persistence of consumption is especially high among the wealthy. Children of the wealthiest half of Danish parents inherit 50% of their parents' excess consumption relative to the mean. These results suggest that studies based on earnings or income data underestimate the intergenerational persistence of living standards.

*Acknowledgement: I thank seminar participants at Aarhus University, the Research Institute of Industrial Economics (IFN), the Swedish Institute for Social Research (SOFI), Stockholm School of Economics, Copenhagen Business School, and the 9th Nordic Summer Institute in Labor Economics for their comments. I also thank the Cycles, Adjustment, and Policy Research Unit (CAP) at Aarhus University and the Knowledge Centre for Housing Economics in Realdania for financial support. I am grateful to the Rockwool Foundation Research Unit for giving me the permission to use their data on Danish neighborhoods. Finally, I thank Henning Bunzel and Søren Leth-Sørensen for helping me assemble the data that I use in this paper.

1 Introduction

A large empirical literature in economics has estimated the extent to which socioeconomic status is transmitted from parents to children. The majority of this work has used data on the earnings or income of parents and children.[1] In this paper I offer a different description of economic mobility by examining the extent to which children inherit the consumption level of their parents. For that purpose I construct a new data sample that merges survey data with registries containing indicators of the living standards of all Danish households.

There are many reasons for why it is important to estimate economic mobility in terms of consumption. A first reason is that consumption is a better indicator of material well-being than earnings or income (Deaton and Zaidi 2002, Meyer and Sullivan 2003). If the ultimate goal is to describe how the living standards of children are determined by their family background, measures of consumption are superior to indirect measures of welfare such as earnings or income.

A second reason is that estimates of the intergenerational persistence of consumption provide valuable evidence on the long-run saving behavior of households. This evidence can be used to test theories of household saving, and to calibrate dynamic models of consumption and wealth inequality. Estimates of the intergenerational persistence of consumption also provide direct evidence on the total consumption smoothing effect of all inter vivos transfers and bequests (realized and anticipated) from parents to children.

One leading theory of household saving is the altruistic household model (Barro 1974, Becker 1974). According to the theory, parents adjust their inter vivos transfers and bequests to smooth their own and their children's consumption.[2] Many studies that feature interactions between parents and their children are based on the altruistic framework (Fahri and Werning 2010, Piketty and Saez 2013). If data on the permanent consumption of parents and children reveal that consumption is no more persistent across generations than earnings or income, such a finding is a strong rejection of the altruistic model and its main behavioral mechanism.

Despite these advantages of using consumption data in mobility studies, there are few previous papers that have estimated the intergenerational persistence of consumption. A first empirical problem is that the typical source of consumption data is a cross-sectional

[1] Papers in this literature include Atkinson, Maynard, and Trinder (1978), Solon (1992), Björklund and Jäntti (1997), Corak and Heisz (1999), Mazumder (2005), and Chetty et al. (2014b).

[2] Examples of previous studies that test this prediction are Altonji, Hayashi, and Kotlikoff (1997) and Behrman and Rosenzweig (2004).

expenditure survey with household data for only one year. What is needed for studies of consumption mobility, however, are measures of household permanent consumption. A second problem is that measures of permanent consumption must be available for both parents and children. Finally, consumption data tends to be noisy, and it is therefore desirable to use an estimation method that allows for measurement error in the permanent consumption of parents and children.

In this paper I combine rich Danish household data and a new identification strategy in a way that I believe goes far towards solving the empirical problems mentioned above. The biggest obstacle when estimating the intergenerational persistence of consumption is measurement error (Solon 1992). I therefore design my estimation strategy to minimize the impact of such errors. First of all, I form the permanent consumption of parents and children with fifteen years of panel data. Secondly, I construct my sample so that it contains parents who participated in the Danish Expenditure Survey. I then use the self-reported consumption of the parents from the survey as an instrument for their permanent consumption as constructed with registry data. Under conditions that I describe below, this method produces consistent estimates of the intergenerational persistence of consumption.

The empirical analysis in this paper produces three main results. The first is that the permanent consumption of Danish households is more persistent across generations than individual earnings and family disposable income. According to the main estimates, the intergenerational elasticity of individual earnings is 0.13, the elasticity of family disposable income is 0.31, and the elasticity of consumption is 0.42. Consumption is thus more than twice as persistent as earnings and roughly 35% more persistent than disposable income. An implication of these results is that estimates of economic mobility based on earnings or income data, overestimate economic mobility in terms of consumption. A country like Denmark, which is known for its high mobility in terms of earnings (Corak 2004), turns out to be considerably less mobile in terms of living standards.

The second result is that the permanent consumption of Danish households regresses to the mean faster than what is predicted by the altruistic household model. The high observed persistence of consumption (relative to earnings and disposable income) is consistent with the altruistic model, but the observed persistence is still lower than the perfect consumption smoothing predicted by the model. Another way of saying the same thing is that the altruistic model exaggerates the extent to which parents smooth their own and their children's consumption.

Finally, the analysis in this paper suggests that there are important differences in the intergenerational persistence of consumption across the parental wealth distribution. In the bottom half of the distribution, the intergenerational elasticity of consumption is approximately 0.3. In the top half of the distribution, the elasticity is as high as 0.5. A full analysis of the mechanisms that generate these differences is beyond the scope of this paper. Nevertheless, it is interesting to note that the high persistence of consumption in the top half of the wealth distribution is consistent with the findings of previous studies showing that wealthy parents have a higher savings rate (Dynan, Skinner, and Zeldes 2004) and leave *proportionally* larger bequests (Menchik and David 1983). The high persistence of consumption among the wealthy is also related to findings in the macroeconomic literature on household saving (Huggett 1996, Krussell and Smith 1998). In one important study, De Nardi (2004) uses a calibrated overlapping-generations model to study the link between household saving and the cross-sectional distribution of wealth. She shows that only with large intended bequests can such a model generate the high concentration of wealth among the very wealthiest families.

A small number of previous papers have estimated the intergenerational persistence of consumption using data from the US Panel Study of Income Dynamics (PSID). In a first study Mulligan (1998) estimates that the intergenerational elasticity of individual earnings in the US is 0.5, the elasticity of family income is 0.6 to 0.7, and the intergenerational elasticity of family consumption is as high as 0.7 to 0.8. The high persistence of consumption relative to earnings and income is consistent with the prediction of the altruistic model that parents smooth their own and their children's consumption. Mulligan does not find any robust differences in the intergenerational persistence of consumption across families with different levels of financial resources.

In a later study, Waldkirch, Ng, and Cox (2004) estimate a structural model of the intergenerational transmission of income and consumption. They find that parents affect the consumption of their children through channels other than their income. More recently, Guo (2014) estimates a switching-regression model to compare the intergenerational persistence of consumption among credit constrained and unconstrained households. Unlike Mulligan, Guo finds that the intergenerational persistence of consumption is highest among constrained (and typically poorer) families. Finally, Charles et al. (2014) estimate the intergenerational persistence of consumption using new and more extensive consumption data in the PSID that has been available since 2005. The authors do *not* find any significant difference between the intergenerational persistence of consumption and income. It is not clear why their results differ from those of Mulligan (1998). Charles et al. work with data of high quality, but one concern is that their measure of permanent consumption is only based on three annual observations of parent and child consumption.

To the best of my knowledge, this paper is the first to estimate the persistence of consumption across generations with data other than the PSID, and thus, the first paper to do so for a country other than the US. The new evidence that is presented in this paper is valuable since previous studies using PSID data have produced conflicting results. There is still no consensus on the extent of intergenerational consumption smoothing in developed economies, and whether the persistence of living standards differs between poor and wealthy segments of the population.

The Danish data also has some strengths relative to the PSID. One advantage is that the Danish expenditure survey is linked to official registries in a way that makes it possible to construct a good instrument for the permanent consumption of parents. Such an instrument is useful given the likely presence of measurement error in consumption. Danish registries also contain high-quality wealth data that is reported by third parties (such as banks) directly to the government. With this data, households can be divided into groups on the basis of their wealth with few classification errors.

In section two below, I present the different types of survey and registry data that are combined in this paper. Section three presents a version of the Becker and Tomes model (1979 and 1986) that is useful when interpreting estimates of the intergenerational persistence of consumption. Section four contains the econometric model and highlights the identifying assumptions in the empirical analysis. Section five presents all the results and section six concludes.

2 Data

2.1 Danish expenditure survey

The main data source I use is the Danish Expenditure Survey which is conducted by Statistics Denmark as an annual, repeated, cross-sectional survey. I begin my sample in 1995 and include each following survey up to 2009 (a total of 15 years of data).

Each household member in the survey is asked to record all his or her private purchases in a diary during a period of two weeks. In addition, the survey includes an interview with the household members in which they are asked to recollect their major purchases during the last 12 months. The household members are also asked about regular expenditures such as their rent, utilities, insurance fees, and transportation costs. Finally, Statistics Denmark imputes a value of owner-occupied housing for households that live in a house or condominium they own.

I use the survey to construct the total self-reported expenditures of a household in a given year. These expenditures are defined as the weighted sum of all purchases of each household member in the two-week diaries (one person in a single household and two people in a married or cohabiting couple), and the purchases recorded during the interview.[3]

2.2 Registry Data

Each member of the households in the Danish Expenditure Survey can be linked to official Danish registries through his or her unique social security number. I can therefore impute a measure of the permanent consumption of all Danish households by linking indicators of consumption in the official registries to the total expenditures of the households in the expenditure survey. This method is similar to those of Skinner (1987) and Blundell, Pistaferri and Preston (2008), who use data from the US Consumer Expenditure Survey to impute a measure of consumption in the PSID. In this section, I describe the different registries that I use for the imputation (more details about the data and the imputation can be found in the online data apendix).

2.2.1 Housing conditions

The first set of registry data contains information on the housing conditions of all Danish households. The main registry in this group is the Buildings and Housing Registry (Bygnings- og boligregistret) which is administered by the Danish government for tax purposes and for the purpose of conducting annual censuses. All housing units in the registry are divided into the three groups of houses, condominiums, and rented apartments. For each unit, the registry contains a list of its residents and a wide range of physical characteristics such as the size of the unit in square meters and the number of rooms.

The housing registry does not contain the value of houses and condominiums, or the rents paid by tenants in rented apartments. I therefore supplement the housing registry

[3] In this entire paper, I treat cohabiting couples as if they were married and henceforth refer to both married and cohabiting couples as married couples.

with three other types of data. The first additional data (Ejendomsvurderingar) is a registry that contains the officicial government appraisals of all houses and condominiums in Denmark (but not rented apartments). The second additional data (Ejendomme salgsoplysninger) is a registry that contains the prices of all houses and condominiums sold in Denmark. The third additional data consists of neighborhoods that were constructed by the Rockwool Foundation to study social segregation in Denmark (Piil Damm and Schultz-Nielsen 2008). The foundation has divided Denmark into 9404 small neighborhoods that are homogenous in terms of their housing conditions, and that are often separated by physical boundaries such as major roads, parks, or water ways. The smaller neighborhoods were also merged into 2296 medium size neighborhoods and 1263 large neighborhoods.[4] In 2004, these neighborhoods had an average of 572, 2343, and 4259 inhabitants respectively.

I estimate three regression models for each year from 1995 to 2009 and use these models to impute the values of houses, condominiums, and rented apartments. In the first model, I regress the sales price of houses on the government appraisal of the house, neighborhood and municipality housing prices, and the physical characteristics of the house. I use this model to impute the value of houses. In the second model, I regress the sales prices of condominiums on the same explanatory variables and use this model to impute the value of condominiums. In the third model, I regress the sales prices of condominiums on all the explanatory variables except for the government appraisal. I use this model to impute the value of rented apartments (for which there is no appraisal).

The fit of the estimated models vary from year to year and the type of housing unit. For most regressions, the set of controls explain close to 80% of the total variation in the market prices of houses and condominiums *within* each of the years from 1995 to 2009.

2.2.2 Car ownership

The second registry I use is the Central Vehicle Registry (Centralregistret for Motorkøretøjer) which is maintained by the Danish government. The registry contains information on the owner, the brand, the model, the type, and the vintage of all private cars in Denmark.[5]

The registry also has information on when a car was first registered in Denmark and when the current owner registered the car. I use the vehicle registry to construct the total

[4] See Hviid et al. (2014) for a description of the largest neighborhoods.

[5] An example of a car in the registry that can help illustrate these terms is a Toyota (brand) Corolla (model) 1.6 (type) made in 1996 (vintage).

number of cars owned by each of the households in my sample in each of the years from 1995 to 2009.

The vehicle registry provides detailed information on car ownership but it does not include the prices of cars. I therefore supplement the registry with data on the prices of new and used cars gathered by the association of Danish car dealers. The price data is structured in a way similar to the vehicle registry with prices of new and used cars classified by their brand, model, type, vintage, and the year in which the car was sold. I match data from the vehicle registry with prices of new and used cars and form a second variable containing the total value of all the cars owned by each of the households in my sample in each of the years from 1995 to 2009.

2.2.3 Dental expenditures

The Danish government provides partial health insurance for the dental care of all Danish residents. The dental treatments that are covered by the insurance plan are delivered by private dentists who are reimbursed by the government for their services. The dentists also charge direct fees to their patients.[6] The third registry that I use (Sygesikringsregistret) contains information on all dental treatments provided to Danish residents through the government health insurance plan. Since the coverage of the plan is (more or less) universal, the registry equivalently contains all dental treatments received by all Danish residents in each year from 1995 to 2009.

The administration of the Danish government health insurance plan is carried out by the Danish regional authorities. At regular intervals, the regional authorities and the Association of Dentists (Tandlægeforeningen) in Denmark agree on a list of prespecified dental treatments, the cost of each treatment, and the share of the cost that should be covered by the government and by the patient.[7]

I combine the information in the health insurance registry with these agreements to form the total expenditures on dental care of each person in my sample in each of the years from 1995 to 2009. I also add up the expenditures of all members of a household to form the total dental expenditures of each household in my sample in each of the years

[6] However, visits to the regular doctor are free in Denmark (the cost is covered by taxes and not by patient fees).

[7] The share of the total cost that is reimbursed by the Danish government varies by treatment and over time. For many treatments, the reimbursement is approximately 40% of the total cost (the patient pays approximately 60% of the cost).

from 1995 to 2009.

One concern may be that a large fraction of dental care is catastrophic in nature and cannot easily be avoided by the patient. I therefore construct two additional variables that also describe each person's dental expenditures. These variables are the total number of routine dental treatments and the total number of teeth cleaning treatments of all members of each household in my sample in each year from 1995 to 2009.

2.2.4 Total expenditures

The final registry that I use is the Income Tax Registry (Indkomstskatteregistret) which is maintained by the Danish tax authorities and contains information on the taxable income of all Danish residents. The registry also contains information on the wealth holdings of all Danish residents, which originates from the period when Denmark had a wealth tax. Due to the tax, banks and other financial institutions had to report the asset holdings of Danish residents directly to the tax authorities (for example their holdings of cash, bonds, and stocks). The institutions were also required to report the debt of all residents (for example their mortgage debt). The wealth tax was abolished in 1998 but the registries still contain updated information on asset holdings and debt. The registry also contains the official government appraisal of each Danish resident's real estate.

Following Browning and Leth-Petersen (2003), I use the Income Tax Registry to form a measure of the total expenditures of the households in my sample in each year from 1995 to 2009. The logic behind the measure is that total household expenditures are equal to household disposable income minus household saving. The Danish Income Tax Registry does not contain information on household saving, but it does contain information on household wealth. If the saving of a household is approximated by the change in household wealth, Δw, one can form a measure of total household expenditures, x, as the difference between household disposable income, y, and the change in household wealth

$$x \approx y - \Delta w \tag{1}$$

To implement this approximation, I use the broadest possible measure of personal income in the tax registry. The measure includes taxable income (for example labor earnings, capital income, payments from government pensions, and payments from unemployment insurance) and many important sources of non-taxable income (for example most non-taxable government transfers). I form the disposable income of a household by adding the income of all household members and subtracting the direct taxes of all members. I form the net wealth of each household in a similar manner by adding the asset holdings of all household members and subtracing their debt. I finally form the change in the wealth of each household by taking the first difference of each household's total net wealth.[8]

2.3 Samples

When I form my data samples, I have to specify what individuals belong to what households. In the Danish Expenditure Survey, the definition of a household is based on information provided by the respondents themselves. In the Danish registry data, the definition of a household is based on information from several different data sources (the housing unit which a group of individuals is registered as living in, the age of the individuals, their potential kinship, and so on). The registry data is the only data that is available for the entire Danish population (and all the individuals that I work with). I therefore restrict my samples to households that can be identified when only registry data is available. The first type of household that I allow for consists of an adult living alone or together with his or her children. The second type of household consists of a couple living alone, or together with their joint and/or individual children. These two household types make up 92% of all households that appear both in the Danish Expenditure Survey and the registry data. For both groups, it is reasonable to assume that the group members form one single economic unit.[9]

From this universe of households, I form three different samples for my analysis. The first sample merges the self-reported consumption of households in the Danish Expenditure Survey with indicators of their consumption from the registry data. I use this sample to impute the consumption of parents and children, and refer to the sample as the "imputation sample". The second sample contains pairs of parents and children, their permanent consumption, and the self-reported consumption of the parents. I use this sample to estimate the intergenerational persistence of consumption, and refer to the sample as the "small mobility sample". The third sample contains pairs of parents and

[8] The imputation method (1) is also used by Leth-Petersen (2010) and Chetty et al. (2014a). The quality of the match between the expenditure data from the Danish household expenditure survey and the expenditure data imputed from official registries (1) is examined by Kreiner, Lassen, and Leth-Petersen (forthcoming). They conclude that "the two measures match each other well on average".

[9] An example of a class of households in the registry data that I discard, are couples living together with an adult who is neither their child nor a parent of one of the spouses. For such groups of individuals, it is difficult to know if the group members form one household and consume goods in common, or if they form two different households with no joint consumption.

children, their after-tax earnings, and their disposable income. The third sample is not derived from the Danish Expenditure Survey and therefore includes a larger number of observations. I use this sample to estimate the intergenerational persistence of after-tax earnings and disposable income, and refer to the sample as the "large mobility sample".

2.3.1 Imputation sample

I form the imputation sample by selecting households that participated in one of the Danish Expenditure Surveys from 1995 to 2009. I require that the head of the household was between 25 and 74 years old when the survey was conducted. I also require that all registry data is available for the head of the household and a potential spouse. These requirements give me an imputation sample with 3546 single households and 5595 married households.

In Table A3 in the online data appendix, I show the correlation matrix for the self-reported consumption of households in the Danish Expenditure Survey and the proxies for their consumption in the registry data. The correlation matrix is based on raw data and does not make any adjustment for measurement error or demographic factors such as age and household composition. As can be seen in the table, all the proxies are *positively* correlated with self-reported consumption. The strongest predictors of household consumption are the value of a household's cars, the value of a household's home, and the number of cars the household owns.

2.3.2 Mobility samples

I form the small mobility sample by selecting pairs of parents and children in which the parent (mother or father) participated in one of the Danish Expenditure Surveys from 1995 to 2009. I require that the parent was at least 40 years old in 1995 and no more than 74 years old in 2009. I require that the child was at least 25 years old in 1995. I also require that all registry data is available for the parent, the child, and potential spouses of the parent and the child, for each year from 1995 to 2009. These requirements give me a small mobility sample with 1737 pairs of parents and children.

I form the large mobility sample with pairs of parents and children, by imposing the same age requirements as for the parents and children in the small mobility sample. However, I do not require that the parent participated in the expenditure survey. I can therefore select the large sample from the entire Danish population. These requirements give me a large mobility sample with 231011 pairs of parents and children.

In Table 1 below, I display the main summary statistics for the small and large mobility samples. The small mobility sample contains slightly more mothers and daughters than fathers and sons. About 80% of the parents and children live in couples. Half of the parents are working, whereas the corresponding number for the children is slighly more than 90%.

Table 1 shows that the parents in the small mobility sample are between 54 and 68 years on average during the sample period, and that their children are between 30 and 44 years on average. These are the ages at which I measure the consumption of parents and children. In principle, I could measure the earnings and the income of parents and children during the same period, but many of the parents have reached an age when they are no longer working. I therefore measure the earnings and income of parents over the ten-year period from 1980 to 1989, when the parents are 39 to 48 years on average.[10] Unfortunately, I cannot construct the consumption of parents for this period because the registries that I work with do not extend that far back in time.

Since I measure the consumption and the earnings and income of parents and children at different stages in the lifecyle, it is important to consider the effect this may have on my estimates. Grawe (2006) points out that households with higher lifetime earnings tend to have steeper age-earnings profiles. If the earnings of parents are measured later in life than the earnings of their children, parents whose measured earnings are very different will have children whose measured earnings are relatively similar. As a consequence, estimates of the intergenerational persistence of lifetime earnings will be biased downwards. In my data, the earnings and income of parents and children are measured at a fairly similar age. The downward bias in my estimates of the persistence of earnings and income is therefore likely to be small.

Studies of lifecycle consumption profiles have shown that consumption tends to track income closely over the lifecycle, and that households with greater lifetime consumption have steeper age-consumption profiles (see for example Carroll 1997 for evidence from the US, and Robb, Magee and Burbidge 1992 for similar evidence from Canada). If the consumption of parents is measured later in life than the consumption of their childen, estimates of the intergenerational persistence of lifetime consumption will also be biased downwards. In my data, the consumption of parents is measured considerably later than the consumption of their children. The downward bias in my estimates of the intergener-

[10] The year 1980 is the first year for which earnings and income data is available in the Danish registries.

ational persistence of lifetime consumption is therefore likely to be bigger.

Overall, I am likely to have a small downward bias in my estimates of the persistence of earnings and income, and a bigger downward bias for consumption. As a consequence, I will *underestimate* the difference between the persistence of consumption and the persistence of earnings and income. The main point that I try to make in this paper is that the persistence of consumption is higher than the persistence of earnings and income. The way in which I construct my sample is thus likely to create a bias that weakens the main result and makes it less significant.

3 Economic Model

This section presents a version of the Becker and Tomes model from Mulligan (1998). It generates the same testable predictions for the intergenerational persistence of consumption as the original models of Becker and Tomes (1979, 1986).[11] The model is useful when interpreting the main empirical results in this paper.

3.1 Budget constraints

The model economy lasts for two periods and consists of parents and their single child. In period one, parents have total resources I^p equal to the sum of their earnings e^p and their inherited wealth a^p. Parents can divide these resources between their own consumption c^p, financial transfers to their child b, and investments in the human capital of their child h. The budget constraint for parents is

$$I^p = e^p + a^p = c^p + b + h \tag{2}$$

In period two, children receive earnings e^c and the financial transfers from their parents with interest $b \cdot (1 + r)$. These resources are used by the children to finance their consumption c^c. The budget constraint for children is

$$e^c + b \cdot (1 + r) = c^c \tag{3}$$

[11] The orginal Becker and Tomes model (1979) is deterministic. The model in Mulligan (1998) includes uncertainty. A slight difference is therefore that the prediction from Mulligan (1998) concerns the stochastic properties of consumption.

The return on the financial transfers from parents to children, $(1+r)$, is equal to the product of an anticipated return, $(1+\bar{r})$, and a random variable λ

$$(1+r) = (1+\bar{r}) \cdot \lambda$$

The random variable λ is distributed according to $\log(\lambda) \sim N(\mu_\lambda, \sigma_\lambda^2)$, and is independent of the total resources of the parents I^p. The parametric assumption for λ generates a closed form equation for the relationship between the consumption of parents and their children. The independence of λ and I^p ensures that all parents have the same opportunities to transfer resources to their children, regardless of their own earnings and wealth.

The earnings of the children in period two, e^c, depend on the human capital investments of parents, h, and the random variable λ, through the earnings function

$$e^c = \lambda \cdot h^v \qquad 0 < v < 1 \tag{4}$$

Investments in human capital are subject to diminishing returns ($v < 1$).

An important consideration is whether parents can borrow freely against the future income of their children (b can be negative). The first Becker and Tomes model from 1979 assumes that they can, whereas the second model from 1986 assumes that they cannot. The second model generates seperate predictions for the intergenerational persistence of earnings and consumption for parents that are and that are not borrowing constrained. Mulligan (1998) shows that these predictions depend on the shape of the earnings function, and Mulligan (1998) and Grawe (2004) therefore argue that it is unwise to use data on the intergenerational persistence of earnings and consumption, to identify groups of families that are borrowing constrained. For this reason, I only test the predictions of the first Becker and Tomes model in which parents can borrow against the future income of their children.[12]

Under this assumption, it is possible to combine the budget constraints of parents (2) and children (3) into one single intergenerational budget constraint for the entire family

$$c^p + \frac{c^c}{1+r} = I^p - h + \frac{e^c}{1+r} \tag{5}$$

The constraint (5) says that the present value of family consumption has to equal the present value of family income net of human capital investments.

[12]This modelling choice reflects what can and cannot be identified from the data I have. It is not related to whether parents in Denmark are borrowing constrained or not.

3.2 Parental preferences

Parents face a tradeoff between how much they consume and how much their child will consume. Parents rank different combinations of their own and their children's consumption according to the function

$$E\left(U\left(c^p, c^c\right)\right) = \frac{\sigma}{\sigma - 1}\left(c^p\right)^{\frac{\sigma-1}{\sigma}} + \theta \cdot E\left(\frac{\sigma}{\sigma - 1}\left(c^c\right)^{\frac{\sigma-1}{\sigma}}\right) \qquad \theta, \sigma > 0 \tag{6}$$

The parameter σ is the elasticity of substitution between parent and child consumption. The parameter θ represents the altruism of parents towards their children. The expectation in (6) is taken with respect to the distribution of the shock λ.

Parents maximize their utility (6) subject to the intergenerational budget constraint (5) and the earnings function (4). Parents can choose their consumption c^p, the financial transfer to their child, b, and the investments in the human capital of their child, h. Parents make their choices prior to the realization of the shock λ, whereas children adjust their consumption after the realization.

3.3 Prediction for consumption

The solution to the parental problem can be used to characterize the intergenerational persistence of consumption (Mulligan 1998)

$$\log\left(c^c\right) = \alpha + \log\left(c^p\right) + \epsilon \tag{7}$$

$$\epsilon = \log\left(\lambda\right) - E\left(\log\left(\lambda\right)\right) \tag{8}$$

As can be seen in equations (7) and (8), the Becker and Tomes model predicts that a regression of the log consumption of children on the log consumption of parents will produce a coefficient of one. Stated differently, the Becker and Tomes model predicts that the consumption of parents and children do not regress to the mean in percentage (or logarithmic) terms. If a son in Denmark has a father who consumes 10% more than his peers, the model predicts that the son will on average consume 10% more than his own peers. A weaker version of this statement is that consumption regresses to the mean slower than earnings and income.[13]

One way to think about the prediction is to view the Becker and Tomes model as an extension of the permanent income model (Friedman 1957). In that model, households

[13] This weaker statement assumes that earnings and income do regress to the mean.

borrow and save to smooth their consumption over the lifecycle. In the Becker and Tomes model, parents adjust the financial transfers to their children and the investments in the human capital of their children, to smooth their own and their children's consumption.

The intercept α in equation (7) depends on the parameters σ, θ, $\bar{r}$, μ_λ, and σ_λ^2. The prediction from the Becker and Tomes model is therefore obtained under the assumption that all parents have the same (isoelastic) preferences, all parents face the same anticipated return on their financial transfers, and all parents face the same uncertainty.[14]

The prediction for the persistence of consumption does not depend on the persistence of earnings and income. If there is one group of families who (for exogenous reasons) have a high persistence of earnings and income and another group of families who have a low persistence, the model predicts that the persistence of consumption will nevertheless be the same in both groups of families (consumption will not regress to the mean in any of the two groups).

Previous research has shown that the persistence of earnings and income is higher in the US than in Denmark (Corak 2004). The Becker and Tomes model predicts that if these differences are caused by factors exogenous to the family, the intergenerational persistence of consumption will be the same in the US and Denmark.[15] The intuition behind this prediction is that parents control the allocation of resources between themselves and their children. Parents can and will therefore neutralize the effects of changes in the economy that would otherwise have an impact on consumption mobility.

4 Econometric Model

The econometric model that I estimate is a regression of the log of the permanent consumption of a child in family i, $\log(C_i^c)$, on the log of the permanent consumption of a parent in family i, $\log(C_i^p)$

$$\log(C_i^c) = \alpha + \beta \log(C_i^p) + D_i\delta + \varepsilon_i \tag{9}$$

[14] In the Becker and Tomes (1979) model with no uncertainty, the condition on the preferences of parents can be replaced with the weaker condition that all parents have identical and homothetic preferences.

[15] This prediction does not require that all families in the US and Denmark have the same preferences, face the same anticipated return on their financial transfers, and face the same uncertainty. It only requires that these parameters are the same for all families *within* each country. The prediction still holds if for example families in the US face more uncertainty.

The model also includes other control variables, D_i. The error term ε_i in (9) has mean zero, and is uncorrelated with the independent variables of the model

$$E(\varepsilon_i) = Cov(\varepsilon_i, \log(C_i^p)) = Cov(\varepsilon_i, D_i) = 0$$

The regression model (9) is the counterpart to equation (7) from the Becker and Tomes model, which describes the relationship between the consumption of parents and children. The regression model allows for an intergenerational elasticity of permanent consumption, β, that is different from one. The estimated value of β can therefore be used to test the prediction that consumption does not regress to the mean across generations.

The consumption of parents and children in the Becker and Tomes model corresponds best to the lifetime consumption of parents and children in Denmark. I do not have access to a sufficiently long data panel to construct these lifetime consumption levels. Instead, I define the log permanent consumption of parents and children as their average log consumption over the fifteen years for which I have data

$$\log(C_i^c) \equiv \left(\frac{1}{T}\right) \cdot \sum_{t=1}^{T} \log(C_{it}^c) \tag{10}$$

$$\log(C_i^p) \equiv \left(\frac{1}{T}\right) \cdot \sum_{t=1}^{T} \log(C_{it}^p) \tag{11}$$

In equations (10) and (11), $\log(C_{it}^c)$ is the log consumption of a child in family i in year t, $\log(C_{it}^p)$ is the log consumption of a parent in family i in year t, and T is the length of the panel (fifteen years).

I construct a measure of the permanent consumption of parents and children by combining information from the Danish expenditure survey with information from Danish registries. This method is described in the next three sections.

4.1 Survey measure of consumption

I denote the self-reported consumption of a household i that participated in the expenditure survey in year t by S_{it}. I assume that the log of self-reported consumption, $\log(S_{it})$, equals the log of true consumption, $\log(C_{it})$, plus a measurement error, e

$$\log(S_{it}) = \log(C_{it}) + e_{it} \tag{12}$$

The measurement error e_{it} has mean zero

$$E(e_{it}) = 0, \tag{13}$$

and is uncorrelated with the the error term in the main regression model (9) and the log of true consumption

$$Cov\left(e_{it}, \varepsilon_i\right) = Cov\left(e_{it}, \log(C_{it})\right) = 0 \tag{14}$$

4.2 Proxy variables for consumption

I assume that the log of true consumption can be written as a linear combination of a vector X of proxy variables and a vector W of demographic variables, plus an imputation error u

$$\log(C_{it}) = X_{it}\zeta_t + W_{it}\gamma_t + u_{it} \tag{15}$$

This specification is similar to that of Skinner (1987), and Blundell, Pistaferri, and Preston (2008). In my data, the proxy variables X are variables from the Danish registries on the housing conditions, car ownership, dental expenditures, and total expenditures of a household.

The imputation error u_{it} in (15) has mean zero

$$E\left(u_{it}\right) = 0 \tag{16}$$

and is uncorrelated with the proxy variables X_{it} and the demographic variables W_{it}

$$Cov\left(u_{it}, X_{it}\right) = Cov\left(u_{it}, W_{it}\right) = 0 \tag{17}$$

The imputation error u_{it} can be correlated across individuals and can be correlated across parents and children.

I also allow for the possibility that the value of a household's home and the total expenditures of a household are measured with error. I model these measurement errors by assuming that the observed levels of these two proxy variables in the registry data, R_{it}, equal the true levels of the proxy variables, X_{it}, plus an error term v

$$R_{it} = X_{it} + v_{it}$$

The error term v_{it} has mean zero

$$E(v_{it}) = 0 \tag{18}$$

and is uncorrelated with the true levels of the two proxy variables X_{it} and the demographic variables W_{it}

$$Cov\left(v_{it}, X_{it}\right) = Cov\left(v_{it}, W_{it}\right) = 0 \tag{19}$$

The error terms v_{it} can be correlated across individuals and can be correlated across parents and children.

4.3 Estimation

According to equation (12), the log of self-reported consumption equals the log of true consumption plus a measurement error. I use this relationship to solve for the log of self-reported consumption as a function of the proxy variables and the demographic variables

$$\log(S_{it}) = X_{it}\zeta_t + W_{it}\gamma_t + u_{it} + e_{it} \tag{20}$$

In a first step, I estimate equation (20) on all observations in the imputation sample.[16] I then impute the log consumption of parents and children in the small mobility sample in each year from 1995 to 2009. This method is equivalent to replacing the log of true consumption, $\log(C_{it})$, with the predicted value of self-reported consumption, $\log(S_{it})$, on the basis of the proxy variables[17], R_{it}, and the demographic variables, W_{it}. In the context of the econometric model, the prediction is equal to

$$P\left(\log\left(S_{it}\right)|R_{it}, W_{it}\right) = \log\left(C_{it}\right) + v_{it}\zeta_t - u_{it} \tag{21}$$

In a second step, I construct the imputed log of the *permanent* consumption of parents and children, $\log\widehat{(C_i^p)}$ and $\log\widehat{(C_i^c)}$, as the average over their imputed log consumption in each sample year

$$\log\widehat{(C_i^p)} \equiv \left(\frac{1}{T}\right) \cdot \sum_{t=1}^{T} P\left(\log\left(S_{it}\right)|R_{it}^p, W_{it}^p\right) \tag{22}$$

$$\log\widehat{(C_i^c)} \equiv \left(\frac{1}{T}\right) \cdot \sum_{t=1}^{T} P\left(\log\left(S_{it}\right)|R_{it}^c, W_{it}^c\right) \tag{23}$$

As can be seen from equation (21), these average are equal to

$$\log\widehat{(C_i^p)} \equiv \left(\frac{1}{T}\right)\left(\sum_{t=1}^{T}\log\left(C_{it}^p\right) + \sum_{t=1}^{T} v_{it}^p\zeta_t - \sum_{t=1}^{T} u_{it}^p\right) \tag{24}$$

$$\log\widehat{(C_i^c)} \equiv \left(\frac{1}{T}\right)\left(\sum_{t=1}^{T}\log\left(C_{it}^c\right) + \sum_{t=1}^{T} v_{it}^c\zeta_t - \sum_{t=1}^{T} u_{it}^c\right) \tag{25}$$

In a third and final step, I estimate the main regression model (9) by replacing the true log of the permanent consumption of parents and children, $\log(C_i^p)$ and $\log(C_i^c)$, with the imputed log of their consumption, $\log\widehat{(C_i^p)}$ and $\log\widehat{(C_i^c)}$.

[16] I instrument for the proxy variables measured with error using variables from the registry data (see the section on estimation below).

[17] To ease the notation, I assume that the proxy variables that are measured without error have a measurement error equal to zero. I can then use the notation R for all proxy variables.

4.4 Main identifying assumptions

Equations (24) and (25) show that the imputed log of permanent consumption is not equal to the true log of permanent consumption. The difference between the two variables is due to the measurement errors in the proxy variables, v_{it}, and the imputation errors, u_{it}. If the main regression model (9) is estimated with ordinary least squares and imputed consumption is used instead of true consumption, the estimated intergenerational persistence of consumption will suffer from attenuation bias. I deal with this problem by estimating the main regression model with two-stage least squares. The instrument I use for the imputed log of the permanent consumption of parents, $\log(\overset{\wedge}{C_i^p})$, is the log of the self-reported consumption of parents from the expenditure survey, $\log\left(S_{it}^p\right)$.

There are two main assumptions that have to be satisfied in order for the 2SLS estimator to be consistent as T goes to infinity (see equations 24 and 25). The first assumption is that the log of the self-reported consumption of parents is uncorrelated with the *mean* measurement error in the proxy variables of parents and children[18]

$$\lim_{T\to\infty} Cov\left(\log\left(S_{it}^p\right), \left(\frac{1}{T}\right)\cdot\sum_{t=1}^{T} v_{it}^k\zeta_t\right) = 0, \quad k = p, c \tag{26}$$

The second assumption is that the log of the self-reported consumption of parents is uncorrelated with the *mean* imputation error of parents and children

$$\lim_{T\to\infty} Cov\left(\log\left(S_{it}^p\right), \left(\frac{1}{T}\right)\cdot\sum_{t=1}^{T} u_{it}^k\right) = 0, \quad k = p, c \tag{27}$$

The first assumption (26) involves two types of data. Self-reported consumption, on the one hand, is the sum of true consumption and a measurement error. The measurement error is primarily due to the length of the expenditure diary (only two weeks) and the mistakes made by the respondents in the survey. The measurement errors in the registry data, on the other hand, are primarily due to mistakes made by the government agencies that maintain the registries. Given the different sources of the data, it is hard so see how the self-reported consumption of parents could be correlated with the measurement errors in the registry data. It is even harder to see how the self-reported consumption of parents

[18]In the errors-in-variables model, measurement error in the dependent variable is not a source of inconsistency. I allow for a nonzero correlation between the measurement errors and imputation errors of parents and children. The two identifying assumptions therefore involve the self-reported consumption of parents and the measurement and imputation errors of both parents and children.

could be correlated with the mean measurement error in the registry data taken over a period of fifteen years. Overall, the first assumption (26) appears to be weak.

The second assumption (27) involves self-reported consumption and the mean of the imputation error. A *sufficient* condition for the assumption to hold, is that the imputation error, u_{it}, exhibits only limited positive serial correlation. In that case, the mean imputation error converges quickly to zero, and the correlation between the log of the self-reported consumption of parents and the mean imputation error will be small.

The second assumption (27) is similar to the assumption that is often made (explicitly or implicitly) in the empirical literature on intergenerational earnings mobility. Many of the papers in that literature construct a measure of the earnings of parents and children by taking an average over several years of earnings observations. The extent to which this method reduces attenuation bias, depends on the amount of positive serial correlation in the measurement error for earnings and the number of observations for each parent (Mazumder 2005). The fact that I have a long panel with as many as fifteen years of observations thus helps minimize the bias in my estimates.

4.5 Empirical specification

I estimate equation (20) seperately for singles and for couples. The total expenditures of a household and the value of a household's home are measured with error. I therefore instrument these two variables with two-year lagged expenditures, mean expenditures by gender and occupation, mean home values by gender and occupation, and mean home values by zip code (a total of four instruments). All the instruments are constructed with registry data.

For singles, the proxies in the registry data, R_{it}, are the total expenditures of the household, the number of cars of the household, the log total value of the cars, two dummy variables for the type of home of the household (a house, a condominium, or a rented apartment), three interaction variables for the type of home and the log value of the household's home, the total dental expenditures of the household, the total number of routine dental treatments of the household, and the total number of teeth cleaning treatments of the household.[19] The demographic variables, W_{it}, are the age of the household head, the number of children living in the household, the age of the youngest and oldest

[19] Some observations contain negative values for total household expenditures. This is the reason why I use total expenditures in levels rather than in logs.

child in the household, three dummies indicating whether the household head is working, retired, and self-employed, and a dummy for the gender of the household head. Since there are relatively few observations on single households in each year, I pool observations for three consecutive years and run a seperate regression for each subsample (with year dummies).

For couples, the proxies in the registry data, R_{it}, are the same as for single households, except that all the variables are the sums of the values for the husband and the wife. The only exception is the housing variables which refer to the housing conditions of the household head (and the spouse). The demographic variables for couples are the same as those for single households, except that both the age of the husband and the wife are included, there is no dummy for the gender of the household head, there is an additional dummy for the couple being legally married, and there are two separate sets of dummies for the husband and the wife working, being retired, and being self-employed. Because there are more observations on couples, I run a seperate regression for couples in each of the years from 1995 to 2009.

Due to economies of scale in consumption, a couple who live together can obtain a higher standard of living than if they had lived apart as singles. For example, a couple can share an apartment whereas two singles have to pay their own rent. I capture this effect by using an equivalence scale for the consumption of couples, which I set equal to the square root of two. This scale is commonly used in applications, and is within the range of estimates in the consumption literature (see Pendakur 1999 for a survey). I then construct the consumption of each spouse by assuming that he or she gets half of the couple's total consumption.

In the main regression model (9), I include a number of additional control variables, D_i. These are the mean age of the parent and the child over the sample period, the mean fraction of years that the parent and child lived in a couple, the mean number of children that lived in the parent and the child household, and the gender of the parent and the child.

The self-reported consumption of parents, S_{it}^p, can come from any of the fifteen years of survey data and from households with different demographic characteristics. I therefore regress the log self-reported consumption of parents on a set of control variables, and use the residual from the regression as my instrument. The control variables in this regression are a dummy for the parent living in a couple, the number of children living in the parental household, the age of the youngest and oldest child in the household, the age of the household head, and year dummies.

5 Results

5.1 OLS results

I first estimate the intergenerational persistence of individual after-tax earnings, family disposable income, and individual consumption with ordinary least squares. I display the results in the upper part of Table 2. Given the likely presence of significant measurement error in these three variables, I include the OLS results mainly as a point of reference for the instrumental variable results that are presented below. To make the results comparable across different specifications and outcomes, I use the same set of control variables D_i in all regressions.

In the entire paper, I report robust asymptotic standard errors for the estimated persistence of earnings and income. For the estimated persistence of consumption, I report bootstrapped standard errors based on 200 draws from the sample. The bootstrapping procedure that I use replicates the whole estimation process, including the imputation of individual housing values, the imputation of log consumption in each sample year from 1995 to 2009, the construction of log permanent consumption, and the estimation of the intergenerational persistence of consumption.

As can be seen in Table 2, the intergenerational elasticity of after-tax earnings is 0.132 and 0.097 when I use the small and the large mobility samples. These results are similar to those of previous studies which have found a low intergenerational persistence of earnings in Denmark. For example, Bratsberg et al. (2007) estimate an intergenerational elasticity of pretax earnings for pairs of Danish fathers and sons that is 0.12.

I then estimate the intergenerational elasticity of family disposable income which is 0.157 and 0.143 when I use the small and the large mobility samples. These estimates for income are slightly higher than the estimates for earnings. The difference between the persistence of individual after-tax earnings and family disposable income is significant at the 1% level when I compare the results from the large mobility sample.

Finally, I estimate the intergenerational elasticity of permanent consumption which is 0.248. This elasticity is higher than the elasticity of both earnings and income. When I compare the estimated persistence of consumption with the estimated persistence of earnings from the small sample, the difference is significant at the 5% level ($p = 0.033$).

When I compare with the large sample, the difference is significant at the 1% level. When I compare the estimated persistence of consumption with the estimated persistence of income from the small mobility sample, the difference is significant at the 10% level ($p = 0.078$). When I instead compare with the large sample, the difference is significant at the 5% level ($p = 0.017$). In terms of magnitudes, the OLS results indicate that consumption is more than twice as persistent as after-tax earnings in Denmark and roughly 60% more persistent than family disposable income.

5.2 IV results

I continue and estimate the intergenerational elasticities of after-tax earnings, family disposable income, and individual consumption with an instrumental variables approach. I present these results in the bottom part of Table 2. The instruments that I use for the earnings and income of parents respectively, are the mean earnings and income in Denmark by gender and three-digit occupational code.

For earnings, the estimated intergenerational elasticities are 0.130 and 0.101 in the small and the large mobility samples. These estimates are not very different from the OLS estimates presented in the top part of Table 2. For family disposable income, the estimated elasticities are 0.313 and 0.283 in the small and large mobility samples. These numbers are higher than the OLS estimates for the persistence of income. The IV estimates for income are also significantly larger than the IV estimates for earnings.

The relatively big difference between the OLS and IV estimates for the persistence of income, may be due to my use of occupation as an instrument for parental income. Solon (1992) shows that if the persistence of income is estimated with an instrument and the instrument is positively correlated with a variable that enters the structural relationship between the income of parents and children, then the IV estimate will be biased upwards. In my application, the occupation of parents may influence the income of children through channels other than parental income. This potential bias works against my main result, since it reduces the difference between the persistence of consumption and the persistence of disposable income.

Finally, I estimate the intergenerational elasticity of consumption using the log of the self-reported consumption of parents as an instrument for their log imputed permanent consumption. In Table 3, I show the results of the first-stage regression in the 2SLS procedure. The coefficient on self-reported consumption is significant at the 1% level and the value of the F-statistic for the excluded instrument is 237.2. These numbers indicate that the identification strategy of this paper does not suffer from a weak-instrument problem.

In the bottom part of Table 2, I show the results from the second-stage regression. The main estimate of the intergenerational elasticity of permanent consumption is 0.422. This elasticity is substantially higher than the elasticities of earnings and disposable income. The difference between the persistence of consumption and the persistence of earnings is significant at the 1% level. The difference between the persistence of consumption and the persistence of income is significant at the 10% level ($p = 0.069$) when the persistence of income is estimated in the large mobility sample. In terms of magnitudes, the persistence of consumption is more than twice as high as the persistence of earnings, and roughly 35% higher than the persistence of disposable income.

To summarize, I obtain qualitative results for Denmark that are similar to those reported by Mulligan (1998) for the US. The estimates for both countries indicate that individual earnings are the least persistent outcome across generations, family disposable income is more persistent, and consumption is the most persistent outcome. The majority of IV estimates are also higher than the corresponding OLS estimates. In terms of absolute numbers, both the OLS and IV estimates show that Denmark has higher intergenerational consumption mobility than the US. This result is a rejection of the prediction from the Becker and Tomes model, that parents use inter vivos transfers and bequests to neutralize differences in earnings and income mobility between groups of families.[20]

5.3 Nonlinearities

In their theoretical analysis, Becker and Tomes (1986) divide families of parents and children into different groups on the basis of parental wealth. Mulligan (1998) and Guo (2014) use a similar classification in their empirical analysis. In a first extension, I follow these authors and estimate the intergenerational persistence of consumption at different levels of parental wealth.

A strength of the Danish data is that the wealth of parents is available for a thirty year period from 1980 to 2010. I classify families according to the wealth of the parents in 1980 (the first observation) and the wealth of the parents in 2010 (the last observation). In

[20] The hypothesis that consumption regresses to the mean across generations at the same rate in Denmark and the US can be rejected at the one percent level.

1980, the parents were 39 years on average and the children 15 years on average. Around that time, the parents made important investments in the human capital of their children. The classification of families based on parental wealth in 1980 is therefore related to the resources that the parents had for such investments. In 2010, the parents were 69 years on average and the children 45 years on average. The wealth of the parents around that time is highly correlated with the wealth that the parents will leave to their children as bequests. The classification of families based on parental wealth in 2010, is therefore related to the amount of financial transfers that the parents will make to their children.

In Table 4, I present the results on the intergenerational persistence of consumption when I divide my sample into two halves on the basis of parental wealth in 1980 and parental wealth in 2010. The top of the table shows the intergenerational elasticity of consumption when the main regression model is estimated with ordinary least squares. When I use parental wealth in 1980 to classify families, the elasticity is 0.236 in the bottom half of the distribution and 0.325 in the top half of the distribution. When I use parental wealth in 2010 for the classification, the elasticity is 0.225 in the bottom half of the distribution and 0.379 in the top half of the distribution.

I continue by estimating the persistence of consumption in the two halves of the parental wealth distribution by using the same 2SLS approach as above. The results are presented in the bottom of Table 4. When I classify families according to parental wealth in 1980, the intergenerational elasticity of consumption is 0.358 in the bottom half of the distribution and 0.471 in the top half of the distribution. When I instead use parental wealth in 2010, the elasticity is 0.345 in the bottom half of the distribution, and as high as 0.588 in the top half of the distribution. Both the OLS and IV estimates thus provide some evidence that the intergenerational persistence of consumption is higher among wealthy families.

5.4 Rank of consumption

All of the results presented so far have been expressed in terms of the intergenerational elasticity of consumption. Another statistic that can be used to summarize intergenerational mobility is the rank correlation for parent and child consumption. Recently, Chetty et al. (2014b) have used the rank correlation for parent and child income to analyze geographic variation in economic mobility in the US. One general advantage of rank correlations over intergenerational elasticities, is that rank correlations are less dependent on outliers. For this reason, I perform a sensitivity analysis in which I estimate the persistence of consumption across generations in Denmark in terms of percentile rank.

The rank correlation for consumption can be obtained through a regression of the percentile rank of child consumption on the percentile rank of parent consumption. The regression coefficient in the model is equal to the rank correlation for parent and child consumption times the ratio of the standard deviations for child and parent consumption. When consumption is measured in terms of rank, the two standard deviations are the same, and the regression coefficient equals the rank correlation.

To implement the linear rank model, I first regress the log permanent consumption of parents and children on the demographic control variables D_i. I then rank the consumption of parents and children on the basis of the residuals from these regressions. Finally, I regress the rank of child consumption on the rank of parent consumption.

In the top half of Table 5, I show the results that I obtain when I estimate the linear rank model with ordinary least squares. For the whole sample, the rank correlation for parent and child consumption is 0.28. This correlation is similar to the intergenerational elasticity of consumption which is 0.24 (when estimated with OLS).

I also estimate rank correlations for parent and child consumption at different levels of parental wealth. When I classify families according to parental wealth in 1980, the rank correlation is 0.239 in the bottom half of the distribution, and 0.334 in the top half of the distribution. When I classify families according to parental wealth in 2010, the rank correlation is 0.225 in the bottom half of the distribution and 0.375 in the top half of the distribution. This last difference is significant at the 10% level ($p = 0.068$).

I continue and estimate the linear rank model with an instrumental variable approach. For that purpose, I regress the log self-reported consumption of parents on the set of demographic control variables D_i. I then create the instrument as the rank of the residual from this regression. In the bottom half of table 5, I show the results from the linear rank model when I estimate the model with 2SLS. In the whole sample, the estimated rank correlation for parent and child consumption is 0.395. This correlation is similar to the estimated intergenerational elasticity of consumption which is 0.422 (when estimated with 2SLS). However, the interpretation is different. The estimated rank correlation implies that a ten percentile point increase in the rank of parent consumption in Denmark, is associated with a four percentile point increase in child consumption on average.

Finally, I estimate the rank correlation for consumption at different levels of parental wealth using 2SLS. When I use parental wealth in 1980 to classify families, the rank correlation is 0.307 in the bottom half of the distribution and 0.459 in the top half of the distribution. When I instead use parental wealth in 2010, the rank correlation is 0.350 in the bottom half of the distribution and 0.424 in the top half of the distribution.

When considered in isolation, each of the regression results that I present offer some evidence that the intergenerational persistence of consumption is higher among wealthy families. One caveat is that only one specification has a difference between the intergenerational elasticity of consumption in the bottom and the top half of the wealth distribution that is statistically significant and only at the 10% level. The joint evidence from all specifications (OLS and IV estimates, consumption measured in logarithms and in terms of rank, at two different points in time as much as thirty years apart) is stronger and reinforces the conclusion that the persistence of consumption is highest among the wealthy. The results based on the rank of consumption are not affected by differences in the variance of parent and child consumption across the parental wealth distribution. It does not seem therefore that the results are purely driven by a higher dispersion of consumption among the wealthy.

5.5 Robustness checks

In this section, I present several robustness checks for the main 2SLS estimates of the intergenerational persistence of consumption. All these results are presented in Table 6.

In the imputation procedure for consumption, I used an equivalence scale to adjust for the difference between single and married households. In a first robustness check, I remove the adjustment so that the consumption of single and married households is measured with the same scale. The estimated intergenerational persistence of consumption is now 0.426 (see column one of Table 6).

Among the proxy variables that I used to impute consumption, were the total expenditures of a household as constructed from the income tax registry. One concern is that I cannot tell from the registry if parents (children) use their expenditures for their own consumption, or if the expenditures include transfers to their children (parents). In column two of Table 6, I show the results from a specification in which I no longer include total expenditures as a proxy for household consumption in the imputation procedure. The new estimate of the intergenerational elasticity of consumption is 0.417. The standard error for the estimate is slightly higher than for the original specification.

Finally, I modify the instrumental variable. Many studies of household consumption are based on nondurable consumption which is less volatile than total consumption. In column three, I show the results from a specification in which I use the self-reported nondurable consumption of parents as an instrument for their total permanent consumption. The new estimate of the intergenerational elasticity of consumption is 0.385.

In the bottom row of Table 6, I turn to robustness checks that are related to the housing conditions of parents and children in my sample. One important concern is that there are differences in housing prices across locations in Denmark. If parents and children live close to each other, these price differences could contribute to the high estimated persistence of consumption for purely mechanical reasons. In column four, I present a robustness check in which I include the log average municipality housing price per square meter, as an additional proxy variable for household consumption. This specification allows individual home values to have a different impact on imputed consumption across municipalities with different housing prices. The intergenerational elasticity of consumption is now estimated to be 0.370.

An additional concern is that the measure of total expenditures from the Danish Expenditure Survey includes an imputed value of owner-occupied housing. This value is likely to be based on the same housing registries that I use to impute household consumption. If there is measurement error in the housing registries, such errors constitute a violation of assumption (26) which states that there can be no correlation between the self-reported consumption of parents (including the value of owner-occupied housing) and the mean measurement error in the proxies for consumption (including the imputed value of a household's home). To avoid this type of bias, I estimate the main regression model with a new instrument which is equal to the self-reported consumption of parents minus their expenditures on housing. The new estimate of the intergenerational elasticity of consumption is 0.451 (see column five of Table 6). Finally, I present my preferred specification in colum six in which I include the average municipality housing price as a proxy for household consumption, and use self-reported consumption minus housing expenditures as the instrument. This last estimate of the intergenerational elasticity of consumption is 0.419.

Overall, the robustness checks in this section have a small impact on the estimated intergenerational elasticity of consumption. The main estimate is robust to changes in the equivalence scale, modifications of the instrumental variable, and the way in which I use data on the housing conditions of parents and children.

6 Conclusion

This paper presents new evidence on the intergenerational persistence of consumption from Denmark. The main estimates indicate that the intergenerational elasticity of consumption among Danish families is approximately 0.4. The rank correlation for parent and child consumption is of a similar magnitude. The persistence of consumption appears to be higher for the wealthy, with an intergenerational elasticity of approximately 0.5 for the top half of families in terms of wealth.

The results in this paper show that family background plays a large role in determining individual living standards in Denmark. The impact of family background is important for the entire population, and especially important for the wealthy. To put the results in some perspective, Denmark is a country with high earnings mobility and substantial government redistribution. The impact of family background on individual living standards is therefore likely to be higher in other countries.

The OLS and IV results both show that there are large and significant differences between the intergenerational persistence of earnings, disposable income, and consumption. Individual earnings are the least persistent outcome across generations, family disposable income is more persistent, and consumption is the most persistent outcome. An important consequence of these results is that estimates of economic mobility based on earnings or income data overestimate mobility in terms of living standards. The results in this paper are similar to those reported for the US by Mulligan (1998), but different from the results in Charles et al. (2014).

The high persistence of consumption in Denmark relative to earnings and disposable income is consistent with the predictions of the Becker and Tomes model. However, the model exaggerates the extent to which parents smooth their own and their children's consumption. For example, the Becker and Tomes model predicts that groups of families with different earnings and income mobility will experience the same consumption mobility. Contrary to this prediction, the US has low earnings and consumption mobility whereas Denmark has higher earnings and consumption mobility.

An open question is why intergenerational consumption mobility is lower than intergenerational earnings and income mobility. The evidence in this paper offers some hints. Admittedly, it is difficult to use intergenerational elasticities and correlations to discriminate between different theories of why consumption is smooth across generations (Piketty 2000). A high intergenerational persistence of consumption is (by itself) consistent with large intended bequests from parents to children, but also consistent with large unintended bequest. More research is therefore needed to understand the behavior of parents and children that induces these mobility patterns.

References

[1] Altonji, Joseph, Fumio Hayashi, and Laurence Kotlikoff (1997). "Parental Altruism and Inter Vivos Transfers: Theory and Evidence." *Journal of Political Economy*, 105, 1121-1166.

[2] Atkinson, Anthony B., Christopher Trinder, and Alan Maynard (1978). "Evidence on Intergenerational Income Mobility in Britain." *Economics Letters*, 1, 383-388.

[3] Barro, Robert J. (1974). "Are Government Bonds Net Wealth?" *Journal of Political Economy*, 82, 1095-1117.

[4] Becker, Gary S. (1974). "A Theory of Social Interactions." *Journal of Political Economy*, 82, 1063-1093.

[5] Becker, Gary S. and Nigel Tomes (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, 87, 1153-1189.

[6] Becker, Gary S. and Nigel Tomes (1986). "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4, S1-39.

[7] Behrman, Jere R. and Mark R. Rosenzweig (2004). "Parental Allocations to Children: New Evidence on Bequest Differences among Siblings." *Review of Economics and Statistics*, 86, 637-640.

[8] Björklund, Anders and Markus Jäntti (1997). "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review*, 87, 1009-1018.

[9] Blundell, Richard, Pistaferri, Luigi, and Ian Preston (2008). "Consumption Inequality and Partial Insurance." *American Economic Review*, 98, 1887-1921.

[10] Bratsberg, Bernt, Røed, Knut, Raaum, Oddbjørn, Naylor, Robin, Jäntti, Markus, Eriksson, Tor, and Eva Österbacka (2007). "Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons." *Economic Journal*, 117, C72-C92.

[11] Browning, Martin and Søren Leth-Petersen (2003). "Imputing consumption from income and wealth information." *Economic Journal*, 113, F282-F301.

[12] Carroll, Christopher D. (1997). "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis." *Quarterly Journal of Economics*, 112, 1-55.

[13] Charles, Kerwin K., Danziger, Sheldon, Li, Geng and Robert Schoeni (2014). "The Intergenerational Correlation of Consumption Expenditures." *American Economic Review Papers and Proceedings*, 104, 136-40.

[14] Chetty, Raj, John N. Friedman, Søren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen (2014a). "Active vs. Passive Decisions and Crowd-out in Retirement Savings Accounts: Evidence from Denmark." *Quarterly Journal of Economics*, 129, 1141-1219.

[15] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014b). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics*, 129, 1553-1623.

[16] Corak, Miles and Andrew Heisz (1999). "The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data." *Journal of Human Resources*, 34, 504-533.

[17] Corak, Miles, ed., *Generational Income Mobility in North America and Europe*. (Cambridge, UK: Cambridge University Press, 2004).

[18] Deaton, Angus and Salman Zaidi (2002). "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." *World Bank Publications*, Number 14101, The World Bank.

[19] De Nardi, Mariacristina (2004). "Wealth Inequality and Intergenerational Links." *Review of Economic Studies*, 71, 743-768.

[20] Dynan, Karen E., Jonathan Skinner and Stephen P. Zeldes (2004). "Do the Rich Save More?" *Journal of Political Economy*, 112, 397-444.

[21] Friedman, Milton, *A Theory of the Consumption Function*, (Cambridge, MA: NBER Books, National Bureau of Economic Research, 1957).

[22] Grawe, Nathan D. (2004). "Reconsidering the Use of Nonlinearities in Intergenerational Earnings Mobility as a Test for Credit Constraints." *Journal of Human Resources*, 39, 813-827.

[23] Grawe, Nathan D. (2006). "Lifecycle bias in estimates of intergenerational earnings persistence." *Labour Economics*, 13, 551-570.

[24] Guo, Sheng (2014). "Switching Regression Estimates of the Intergenerational Persistence of Consumption." *Economic Inquiry*, 52, 1503-1524.

[25] Huggett, Mark (1996). "Wealth distribution in life-cycle economies." *Journal of Monetary Economics*, 38, 469–494.

[26] Hviid, Simon, Christopher D. Timmins, and Rune Vejlin (2014). "A Dynamic Model of Residential Mobility: Valuation of Non-Marketed Local Amenities." Working Paper., Aarhus University.

[27] Kreiner, Claus Thustup, Lassen, David Dreyer , and Søren Leth-Petersen, "Measuring the Accuracy of Survey Responses using Administrative Register Data: Evidence from Denmark." Forthcoming in *Improving the Measurement of Household Consumption Expenditures*, edited by Christopher Carroll, Thomas Crossley, John Sabelhaus. (NBER Book Series Studies in Income and Wealth)

[28] Krusell, Per and Anthony A. Smith (1998). "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106, 867-896.

[29] Leth-Petersen, Søren (2010). "Intertemporal Consumption and Credit Constraints: Does Total Expenditure Respond to an Exogenous Shock to Credit?" *American Economic Review*, 100, 1080-1103.

[30] Mazumder, Bhashkar (2005). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Review of Economics and Statistics*, 87, 235-255.

[31] Menchik, Paul and Martin David (1983). "Income Distribution, Lifetime Savings, and Bequests." *American Economic Review*, 73, 672-690.

[32] Meyer, Bruce and James Sullivan (2003). "Measuring the Well-Being of the Poor Using Income and Consumption." *Journal of Human Resources*, 38(S), 1180-1220.

[33] Mulligan, Casey, *Parental Priorities and Economic Inequality*, (Chicago, IL: University of Chicago Press, 1998)

[34] Pendakur, K. (1999). "Semiparametric Estimates and Tests of Base-Independent Equivalence Scales." *Journal of Econometrics*, 88, 1-40.

[35] Piil Damm, Anna and Marie Louise Schultz-Nielsen (2008). "Danish Neighbourhoods: Construction and Relevance for Measurement of Residential Segregation." *Nationaløkonomisk Tidsskrift*, 146, 241-262.

[36] Piketty, Thomas, Theories of persistent inequality and intergenerational mobility, in *Handbook of Income Distribution*, eds. Anthony. B. Atkinson and Francois Bourguignon, Edition 1, Volume 1, Chapter 8, 429-476, (Elsevier 2000).

[37] Robb, A. Leslie, Lonnie Magee, and John B. Burbidge (1992). "Kernel Smoothed Consumption-Age Quantiles." *Canadian Journal of Economics*, 25, 669-80.

[38] Skinner, Jonathan (1987). "A superior measure of consumption from the panel study of income dynamics." *Economics Letters*, 23, 213-216.

[39] Solon, Gary (1992). "Intergenerational Income Mobility in the United States." *American Economic Review*, 82, 393-408.

[40] Waldkirch, Andreas, Serena Ng, and Donald Cox (2004). "Intergenerational Linkages in Consumption Behavior." *Journal of Human Resources*, 39, 355-381.

TABLE 1

Summary statistics for the mobility samples

Small Mobility Sample		
	Parents	Children
Total number of observations	1737	1737
Total number of men	729	761
Total number of women	1008	976
Mean age consumption	54-68	30-44
Mean age earnings/income	39-48	30-44
Fraction couples 1995-2009	0.799	0.814
Fraction working 1995-2009	0.493	0.912
Fraction retired 1995-2009	0.426	0.022

Large Mobility Sample		
	Parents	Children
Total number of observations	231011	231011
Total number of men	105600	107662
Total number of women	125411	123349
Mean age earnings/income	39-48	29-43
Fraction couples 1995-2009	0.783	0.792
Fraction working 1995-2009	0.494	0.911
Fraction retired 1995-2009	0.404	0.027

Note: This table displays summary statistics for the samples that are used to estimate the intergenerational persistence of individual earnings, family disposable income, and consumption in Denmark.

TABLE 2

Estimates of economic mobility

OLS

	After-tax Earnings		Family Disposable Income		Consumption
Sample	Small	Large	Small	Large	Small
β	0.132***	0.097***	0.157***	0.143***	0.248***
	(0.032)	(0.0025)	(0.027)	(0.0033)	(0.044)
R2	0.073	0.065	0.570	0.575	0.796
# observations	1716	231011	1716	231011	1737

2SLS

	After-tax Earnings		Family Disposable Income		Consumption
Sample	Small	Large	Small	Large	Small
β	0.130***	0.101***	0.313***	0.283***	0.422***
	(0.035)	(0.0027)	(0.055)	(0.0073)	(0.076)
# observations	1690	231011	1690	231011	1737

Note: Each column refers to a different regression. Dependent variable in all regressions is log child outcome. Unit of observation is pair of parent and child. For earnings and income, numbers in parentheses are robust asymptotic standard errors. For consumption, numbers in parentheses are bootstrapped standard errors. All regressions include additional demographic control variables D_i that are listed in section 4.5. *** Significant at 1%.

TABLE 3

First-stage regression

Log self-reported expenditures	0.180***
	(0.020)
# observations	1737
R2	0.746
F-test excluded instrument	237.2

Note: Dependent variable is log permanent parent consumption. Unit of observation is parent household. Numbers in parentheses are bootstrapped standard errors. The regression includes additional demographic control variables D_i that are listed in section 4.5. *** Significant at 1%.

TABLE 4

Estimates of consumption mobility by parental wealth

OLS				
	Parental Wealth 1980		Parental Wealth 2010	
Sample	Bottom 50%	Top 50%	Bottom 50%	Top 50%
β	0.236***	0.325***	0.225***	0.379***
	(0.066)	(0.080)	(0.064)	(0.096)
R2	0.808	0.746	0.787	0.793
# observations	623	624	530	531

2SLS				
	Parental Wealth 1980		Parental Wealth 2010	
Sample	Bottom 50%	Top 50%	Bottom 50%	Top 50%
β	0.358***	0.471**	0.345**	0.588**
	(0.10)	(0.22)	(0.16)	(0.26)
# observations	623	624	530	531

Note: Each column refers to a different regression. Dependent variable in all regressions is log child permanent consumption. Unit of observation is pair of parent and child. Numbers in parentheses are bootstrapped standard errors. All regressions include additional demographic control variables D_i that are listed in section 4.5. *** Significant at 1%, ** significant at 5%.

TABLE 5

Estimates of consumption mobility in terms of rank

OLS

	All Parents	Parental Wealth 1980		Parental Wealth 2010	
Sample	Small	Bottom 50%	Top 50%	Bottom 50%	Top 50%
β	0.280***	0.239***	0.334***	0.225***	0.375***
	(0.035)	(0.056)	(0.055)	(0.058)	(0.058)
R2	0.079	0.056	0.11	0.051	0.14
# observations	1737	623	624	530	531

2SLS

	All Parents	Parental Wealth 1980		Parental Wealth 2010	
Sample	Small	Bottom 50%	Top 50%	Bottom 50%	Top 50%
β	0.395***	0.307***	0.459***	0.350***	0.424***
	(0.071)	(0.11)	(0.16)	(0.14)	(0.17)
# observations	1737	623	624	530	531

Note: Each column refers to a different regression. Dependent variable in all regressions is rank of child permanent consumption. Unit of observation is pair of parent and child. Numbers in parentheses are bootstrapped standard errors. All regressions include additional demographic control variables D_i that are listed in section 4.5. *** Significant at 1%.

TABLE 6

Robustness checks for estimates of the persistence of consumption

2SLS			
	(1)	(2)	(3)
β	0.426***	0.417***	0.385***
	(0.076)	(0.11)	(0.077)
# observations	1737	1767	1737

2SLS			
	(4)	(5)	(6)
β	0.370***	0.451***	0.419***
	(0.081)	(0.089)	(0.095)
# observations	1737	1737	1737

Note: Each column refers to a different regression. Dependent variable in all regressions is log child permanent consumption. Unit of observation is pair of parent and child. Numbers in parentheses are bootstrapped standard errors. The different columns contain the following robustness checks: (1) no equivalence scale for single and married households, (2) total expenditures from tax registry not used in imputation, (3) instrument is log self-reported nondurable consumption, (4) log average municipality housing price per square meter is additional proxy variable for consumption in imputation, (5) instrument is log self-reported consumption minus housing expenditures, (6) log average municipality housing price per square meter is additional proxy variable for consumption in imputation and instrument is log self-reported consumption minus housing expenditures (combination of checks 4 and 5). All regressions include additional demographic control variables D_i that are listed in section 4.5. *** Significant at 1%.

Data Appendix (Only for online publication)

Danish Expenditure Survey

The basis for the data sample is the Danish Expenditure Survey which is conducted by Statistics Denmark. The latest version of the survey was initiated in 1994, but data from that year cannot be linked to registry data. I therefore use each annual survey from 1995 up to 2009 (a total of 15 years). The expenditure survey is a repeated cross-section that covers the entire noninstitutionalized Danish population above age 15. A total of 1500 households are drawn for the survey annually, but slightly less than half of those households complete the entire survey. In Table A1, I show the number of singles and couples (married or cohabiting) from the expenditure survey that are included in my imputation sample.

As part of the Danish Expenditure Survey, each household member is requested to record all his or her private purchases in a diary during a period of two weeks. The household members record a detailed description of the goods and services that they purchase, and the purchase price of each good or service. The expenditure survey also contains a longer interview with the household members that is conducted in their home. During this interview, the household members are asked to recollect major purchases they made during the last 12 months. The household members are also asked about regular expenditure items such as their rent, utility costs, insurance fees, and transportation costs. Finally, Statistics Denmark imputes a value of owner-occupied housing (for households that live in a condominium or a house they own).

The responses from the households are coded into an expenditure system with approximately 1200 different categories of goods and services. I form the total expenditures of each household by adding up all the purchases from the diaries (after scaling the two-week period into an entire year) and the purchases recorded in the interview. For each household, I also form the total expenditures on nondurable goods and services, and total expenditures minus expenditures on housing.

Imputation of home values

I first combine data from the Buildings and Housing Registry, data on Danish neighborhoods (from the Rockwool foundation) and municipalities, and the prices of sold houses and condominiums in Denmark, to compute average local housing prices in each year from

1995 to 2009. More specifically, I compute the average price of houses and condominiums per square meter and per room at the neighborhood and municipality level. I then form four different predicted market prices for each housing unit in my sample, by multiplying the average price per square meter and per room at the neighborhood and municipality level, with the actual square meters and number of rooms of the housing unit.

I compute the average neighborhood price for the smallest neighborhood (small, medium, or large) in which 25 or more houses or condominiums were sold in a year. For 1% of the housing units, the neighborhood price data comes from the small neighborhoods, for 27% it comes from the medium size neighborhoods, and for an additional 46% it comes from the large neighborhoods. For the last 26% of the housing units, there are not enough sales in the large neighborhood, and I therefore use average prices at the municipality level.

I then impute the values of houses, condominiums, and rented apartments with three different regression models. I estimate three separate models in each year form 1995 to 2009.

In the first model, I regress the sales price of houses on the official government appraisal of the house, the size of the house in square meters, the number of rooms in the house, the year the house was constructed, the number of restrooms in the house, and the four different predicted market prices of the house (on the basis of size in square meters, number of rooms, and the average prices in the neighborhood and the municipality per square meter and per room). I use this model to impute the value of the homes of those households who live in houses.

In the second model, I regress the sales price of condominiums on the official government appraisal of the condominium, and the same set of control variables that I use for houses. I use this model to impute the value of the home of those households who live in condominiums. Finally, I regress the sales price of condominiums on the same set of control variables except for the official government appraisal of the condominium. I use this model to impute the value of the home of those households who live in rented apartments (for which there is no official government appraisal).

Car data

The data on the ownership of private cars comes from the Central Vehicle Registry (Centralregistret for Motorkøretøjer) which is maintaineed by the Danish government. A small number of individuals are registered as owning a large number of cars. I suspect that these are collectors of cars and/or car dealers. I therefore cap the number of cars that each person can own to three cars. A couple can own a total of six cars (three cars for each spouse).

The prices of new and used cars are based on the reports of individual car dealers in Denmark who submit information on their car sales to the central office of the Danish Car Dealer's Association (DAF). The prices that I use are formed as averages across all these reported sales. The administration of the price data is carried out by the association's subsidiary company AutoIT.

The data on the prices of new and used cars is structured in a way similar to the Central Vehicle Registry. Unfortunately, there is no formal identification code in the price data that makes it possible to obtain a perfect match with the vehicle registry. I therefore had to merge the two data sets manually. For 25% of all cars in my sample, the match is based on all relevant parameters (brand, model, type, vintage, and year). For another 61%, the match is based on brand, model, vintage and year (by taking an average across types). For the rest of the matches, I also took an average across models, and (in very rare cases) across brands.

Some Danish households have access to company cars that they can use outside of work, but which are owned by their employer. This arrangement is most common for managers in large private firms. I cannot identify these cars in the vehicle registry since they are registered with the employer. However, individuals who have access to company cars have to pay a tax based on the fringe benefit they receive. Since the size of the tax depends on the value of the car, I can use the official Danish tax registries to identify company cars and approximate their value.

The information about company cars is more noisy than the information from the vehicle registry, since the tax on fringe benefits is based on the total value of all fringe benefits. According to Statistics Denmark, 70% of all taxable fringe benefits are due to company cars. This fraction is likely to be higher for individuals who have access to company cars and therefore receive large taxable fringe benefits.

In Table A2, I show the main features of the tax rules for cars that are up to three years old. These are the rules that I use to convert data on taxable fringe benefits into information on who has access to a company car, and the market value of this car. As can be seen in the table, the tax rules can be summarized in terms of three parameters: the rate at which the value of a car is converted to a taxable fringe benefit, the minimum value to which the rate is applied (the minimum value of the car), and the maximum value to which the rate is applied (the maximum value of the car). Up until year 2000, the official tax rules contained all three of these parameters. After year 2000, there is no longer a maximum value for cars. In order to avoid extreme outliers, I nevertheless set the maximum value of a car to 450000 Danish crowns for the years 2001 to 2009.

I assign each person in my sample (a single person or any of the two spouses in a couple) zero or one company car. By default, a person has zero company cars. I assign one company car to the person if he or she has taxable fringe benefits that exceed the minimum benefit for a company car. For example, if a person had taxable fringe benefits of at least 30000 Danish crowns in 1995, I assign him or her one company car in that year. For those individuals who are assigned a company car, I invert the tax rules to find the market value of the car. For example, if a person had taxable fringe benefits of 50000 Danish crowns in year 2000, I assign him or her one company car in that year with a value of 200000 Danish crowns.

Self-reported consumption and proxies for consumption

In Table A3, I show the correlation matrix for the self-reported consumption of the households that participated in the Danish Expenditure Survey, and the proxies for their consumption in the registry data. The correlation matrix is based on raw data and does not make any adjustments for measurement error or demographic factors such as household age and composition. As can be seen in the table, all of the proxies that I use are positively correlated with self-reported consumption.

Parental wealth

I divide the small mobility sample into two halves on the basis of the total wealth of a child's two biological parents. This division requires that I have wealth information for both of the biological parents. Due to missing data for some of these parents, the subsamples based on parental wealth are slighlty smaller than (half of the) small mobility sample.

TABLE A1

Number of observations in the imputation sample

Year	Singles	Couples
1995	238	429
1996	271	428
1997	263	400
1998	273	385
1999	262	408
2000	286	395
2001	242	412
2002	153	241
2003	247	349
2004	221	353
2005	175	241
2006	275	412
2007	204	312
2008	225	388
2009	211	442
Total	3546	5595

Note: This table displays the total number of observations (the total number of singles and couples) in the Danish expenditure survey that can be linked to a full set of registry variables.

TABLE A2

Taxable value of company cars

Year	Rate	Minimum Value	Maximum Value
1995	20%	150000	400000
1996	21%	150000	400000
1997	22%	150000	400000
1998	23%	150000	400000
1999	24%	160000	450000
2000	25%	160000	450000
2001	25% up to 300000, 20% above	160000	none
2002	25% up to 300000, 20% above	160000	none
2003	25% up to 300000, 20% above	160000	none
2004	25% up to 300000, 20% above	160000	none
2005	25% up to 300000, 20% above	160000	none
2006	25% up to 300000, 20% above	160000	none
2007	25% up to 300000, 20% above	160000	none
2008	25% up to 300000, 20% above	160000	none
2009	25% up to 300000, 20% above	160000	none

Note: This table shows the official tax rules that are used to convert the market value of a company car into a taxable value of the car for the employee (for cars that are no more than three years old). The rate is the fraction of the market value that is used to generate the taxable value, and the miminum and maximum values are the minimum and maximum market values of a car to which the rate applies.

TABLE A3

Correlation matrix for self-reported consumption and proxy variables

Couples (n = 16344)

	S_{it}	Cars	Price	Home	Dental	Routine	Cleaning	$y_{it} - \Delta w_{it}$
S_{it}	1.000	0.183	0.322	0.240	0.041	0.105	0.064	0.137
Cars	0.183	1.000	0.579	0.040	0.029	0.056	0.041	0.045
Price	0.322	0.579	1.000	0.127	0.030	0.102	0.067	0.091
Home	0.240	0.040	0.127	1.000	0.040	0.089	0.037	0.006
Dental	0.041	0.029	0.030	0.040	1.000	0.570	0.368	-0.005
Routine	0.105	0.056	0.102	0.089	0.570	1.000	0.613	-0.002
Cleaning	0.064	0.041	0.067	0.037	0.368	0.613	1.000	-0.011
$y_{it} - \Delta w_{it}$	0.137	0.045	0.091	0.006	-0.005	-0.002	-0.011	1.000

Singles (n = 10470)

	S_{it}	Cars	Price	Home	Dental	Routine	Cleaning	$y_{it} - \Delta w_{it}$
S_{it}	1.000	0.302	0.386	0.139	0.060	0.105	0.090	0.097
Cars	0.302	1.000	0.664	0.060	0.052	0.113	0.093	0.055
Price	0.386	0.664	1.000	0.076	0.047	0.096	0.072	0.073
Home	0.139	0.060	0.076	1.000	0.030	0.046	0.034	-0.021
Dental	0.060	0.052	0.047	0.030	1.000	0.546	0.347	-0.008
Routine	0.105	0.113	0.096	0.046	0.546	1.000	0.602	-0.009
Cleaning	0.090	0.093	0.072	0.034	0.347	0.602	1.000	-0.010
$y_{it} - \Delta w_{it}$	0.097	0.055	0.073	-0.021	-0.008	-0.009	-0.010	1.000

Note: Correlation matrix for self-reported consumption, S_{it}, and proxies in registry data. Each correlation is weighted average over correlations in each sample year 1995-2009. Proxy variables are number of cars (cars), total price of all cars (price), value of home (home), total dental expenditures (dental), number of routine dental treatments (routine), number of cleaning dental treatments (cleaning), and total expenditures from tax registries ($y_{it} - \Delta w_{it}$).