UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THREE ESSAYS IN DEVELOPMENT AND GROWTH

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

WEICI YUAN
Norman, Oklahoma
2016

THREE ESSAYS IN DEVELOPMENT AND GROWTH

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF ECONOMICS

BY

____
Dr. Kevin Grier, Chair

____
Dr. Scott Linn

____
Dr. Robin Grier

____
Dr. Gregory Burge

____
Dr. Daniel Hicks

Acknowledgements

I would like to express my sincere gratitude to my committee members: Kevin Grier, Daniel Hicks, Robin Grier, Gregory Burge, and Scott Linn. This dissertation would not have been possible without their generous help. Special thanks to Kevin Grier and Daniel Hicks who have shown me how to become a qualified scholar. Finally, I want to thank my parents Jianren Li and Mingmin Yuan, and my wife Huiqiong Duan. They have always been supportive through the whole process.

Table of Contents

List of Tables

List of Figures

Abstract

My research explores important issues associated with development and growth. Chapter 1 aims to obtain an accurate estimate of China's intergenerational income mobility and to present evidence on its distributional pattern. Using panel data from the China Health and Nutrition Survey (CHNS) over the period 1989—2009, I find that China is less mobile than most developed countries. Then, I employ five different approaches to investigate the distributional pattern of China's intergenerational mobility across income levels. The results suggest that poor families have relatively high mobility, indicating opportunities for the poor children to escape poverty. Finally, I show that while wealthy fathers are likely to pass on their favorable economic status to their sons, rich sons come from a very wide range of family economic backgrounds.

In the second chapter of my dissertation, conspicuous consumption in China is studied. Conspicuous consumption, also called visible expenditures, conveys higher socioeconomic status and may help individuals differentiate themselves in the marriage market when there is competition for partners and imperfect information. I examine a unique dataset of automobile purchasers in China to investigate the extent to which skewed sex ratios influence expenditure decisions for this highly visible commodity. Using a triple difference approach, I show that unmarried male consumers who face an unfavorable sex ratio purchase more expensive, luxury vehicles than their married peers. Lower income borrowers and those residing in regions with the worst sex ratios exhibit the largest relative degree of conspicuous consumption. In addition to the direct cost of consumption signaling, I demonstrate this behavior generates negative externalities in the form of lower average fuel economy and higher average vehicle weight. As it has

worsened sex ratios, status competition and the associated negative repercussions I identify represent unintended consequences of China's one-child policy.

Chapter 3 evaluates the causal effect of equity market liberalization in sixty-five countries that have adopted this policy during 1980-2011. While previous research has been devoted to liberalization cases prior to 2000, I extend the data through 2011, which leads to a roughly 50% increase in the number of cases. I first replicate the results in Bekaert et al. (2005) and then extend their analysis to the present. I find that the previously estimated growth effects remain significant in the updated sample. Next, I seek to address issues of endogenous policy selection that existing literature has failed to handle adequately. To do this, I model the liberalization decisions using a Cox proportional hazard regression and combine propensity score matching with difference-in-difference methods to obtain an unbiased estimate of the treatment effects of liberalization. My results suggest that equity market liberalizations generate growth effects that are far more persistent than previously documented.

Chapter 1: The Sins of the Fathers: Intergenerational Income Mobility in China

1.1. Introduction and Literature

Income inequality has always been a major concern for both economists and politicians. One way to measure inequality is to examine the income distribution at a given point in time, typically using cross-sectional data. On the other hand, intergenerational income mobility deals with the way current inequality is passed to the next generation.

There is a large body of literature exploring intergenerational mobility, much of which focuses on fathers and sons in the US. The conventional approach is to estimate the intergenerational income elasticity (IGE) which gives an answer to the question: If the father's lifetime income is one percent higher than the average of his generation, by how many percentage points will his son's lifetime income exceed the average of the second generation. The IGE is a mirror image of the intergenerational income mobility. They are inversely related to each other. In the following paper, both of these terms will be used.

Most of the early papers find the IGE in the USA to be about 0.2. However, in later research, people point out that the single-year measure of father's earnings induces an attenuation bias given that it is a poor proxy for the permanent income. As a result, they use better data such as Panel Study of Income Dynamics (Solon, 1992) and National Longitudinal Surveys of Labor Market Experience (Zimmerman, 1992) and replace the single-year father's earnings with averages of father's earnings taken over three to five years. They conclude that the IGE in the USA over the long run is around 0.4. More recently, Mazumder (2005) argues that even 0.4 has been biased down by 30% due to the persistent transitory income fluctuations and that the USA is substantially less mobile than people have thought.

Compared with the research on the USA, many studies find higher mobility in other OECD countries such as Sweden, Germany, Finland, Norway, Spain, and Canada.[1] But there are two exceptions: Great Britain's IGE almost reaches 0.6 when IV regression is employed (Dearden et al., 1997); and the IGE in Italy is estimated to be 0.55 or 0.44, depending on the definition of income used (Piraino, 2007).

However, only a very limited number of papers focus on the less developed countries due to the paucity of data. According to these studies, mobility appears to be lower on average in developing countries such as Brazil, Ecuador and Peru (Andrade et al., 2004; Grawe, 2004; Dunn, 2007; Gong et al., 2012). Table 1.1 provides the main findings of relevant papers.

In recent years, there have been a number of works regarding China's intergenerational mobility, but as of yet there is no consensus. The estimates of China's IGE range from around 0.3 to 0.63 (Guo and Min, 2008; Gong et al., 2012; Fan et al., 2013; Yuan and Lin, 2013). The broad range of IGE estimates probably result from different samples, sample selection rules, econometric model specifications and definitions of income.

Aside from obtaining estimates of the IGE in each country, researchers are also interested in the distributional pattern of intergenerational mobility across income levels. However, the literature on this topic is still very small. To my knowledge, only five

[1]See Bjorklund and Jantti, 1997; Couch and Dunn, 1997; Corak and Heisz, 1999; Osterbacka, 2001; Bratburg et al., 2005 and Pascual, 2009.

papers carefully examine the issue econometrically.[2] In these studies, Eide and Showalter (1999) and Andrade et al. (2004) perform quantile regressions. Interestingly, The USA is found to have a generally decreasing IGE while Brazil is almost the opposite. Corak and Heisz (1999) use a nonparametric model and find that income mobility in Canada is higher at the lower end of the income distribution than on the top and in the middle. Finally, Bjorklude et al. (2012) use linear spline regressions to show that for the 0.1-percent richest Swedish families, fathers' economic status is highly transmissible to their sons. Figure 1.1 summarizes these findings.[3]

My paper is the first one to study in depth how intergenerational mobility varies across income levels in China. Using a diversity of methods, I find that poor families enjoy higher mobility, which may give people more confidence in China’s poverty reduction. On the other hand, whereas wealthy fathers tend to give rise to wealthy sons, wealthy sons can come from a broad range of family economic backgrounds.

The remainder of this paper will be organized as follows: Section 2 introduces the data and sample selection rules, Section 3 estimates China's overall IGE, Section 4 investigates its distributional pattern, and Section 5 concludes the paper.

1.2. Data and Sample Selection

The data are from the China Health and Nutrition Survey (CHNS), which was conducted by the Carolina Population Center at the University of North Carolina at

[2] Although there are many papers approaching the distributional pattern problem using transition matrices, they all reach similar conclusions due to an innate flaw of the matrix. I will discuss and correct the flaw in section 4.

[3] Grawe (2004) applies two-sample-two-stage-least-squares (TS2SLS) quantile regressions and estimated the IGE distribution patterns for several countries. Since most of their samples are quite small, I do not show the graphs.

Chapel Hill and the Chinese National Institute of Nutrition and Food Safety. The survey covers a total of nine provinces in China including Heilongjiang, Jiangsu, Shandong Liaoning, Henan, Hubei, Hunan, Guangxi and Guizhou, and has had eight waves, collected in 1989, 1991, 1993, 1997, 2000, 2004, 2006 and 2009 respectively. In each wave of the survey, approximately 200 communities, 4000 households, and 26000 individuals were interviewed. The participants were asked about health and nutrition conditions, medical care, family planning, demography, education, socioeconomic status, among other things. This micro longitudinal data set has proven to be representative and reliable (Wang, 2007), and is regarded as one of the best resources to investigate Chinese households and communities.

There are advantages and limitations regarding the data. The advantages are: first, it is a longitudinal data set that spans twenty years. If a person takes the survey in multiple years, researchers will be able to have a better knowledge of that person's income trend and lifetime income. Second, it collects information on individuals whether or not they are still residing in the same household with their families. Therefore, it does not suffer from the co-residing bias. The biggest shortcoming of the data, however, is that not every household participates in the survey from 1989 to 2009. Households may exit the survey for reasons that researchers may not know although the data do tell researchers when an individual leaves the sample due to death. New households will enter to replace the old ones so that the total number of families and individuals interviewed is similar in each wave. Thus, the sample includes households whose survey years may be different. To address this issue, I only use the households that stay in the survey for more than 16 years so that they are from roughly the same period.

I utilize the family member relationship file and individual's gender information to identify all the father-son cases. Sons who are enrolled in school and fathers who have retired are excluded from the sample. When calculating permanent incomes, years with negative or zero incomes are not considered. This step is justified by more than a computational issue. Reports of abnormally low incomes are most likely to be a result of measurement errors and using these values could incorrectly assign the corresponding individuals a very low lifetime income. The problems regarding missing values are always salient when dealing with survey data. In this study, I do not exclude individuals who fail to have a complete income series. For example, if one out of eight years is missing when calculating the eight-wave average, I use the average of the remaining seven years. This exercise is consistent with Osterberg (2000) and Bratberg (2005). However, I drop fathers who have fewer than five income observations to ensure that the average incomes can approximate the lifetime incomes. Finally, following Couch and Dunn (1997) and Mazumder (2005), if more than one son is matched to a father, all sons who satisfy the screening rules are retained to have a larger sample size.[4] After all the restrictions are imposed on the data, the final sample size is reduced to 1407 father-son pairs, which is fairly small compared with the raw data. I acknowledge that there might be a representativeness issue associated with the small final sample used in the regressions. Table 1.2 reports the summary statistics of the key variables. Here, the concept "Income" is defined as incomes from all sources including job earnings (calculated as the product of monthly earnings and the number of months a person worked in a given year), annual bonus and other cash or non-cash incomes. It is

[4]The standard errors are adjusted for within-household correlation. If the sample is restricted to the oldest son in the household, there are fewer observations, but the results are largely unchanged.

then adjusted to the 2009 price level using the consumer price indices. It shows that the sons earn more than the fathers on average. Sons are much better educated than fathers, probably due to the 9-year compulsory education since 1986 and the increasing return on human capital in the last decades (Yuan and Lin, 2013). The average age is about 24 for the sons and 53 for the fathers in 2000.

1.3. China's Overall Mobility

1.3.1. Empirical Model

I use a Galton-Becker-Solon equation as the baseline regression model. It is a conventional specification in the literature (Solon, 1992; Solon, 2002):

(1) $$Y_i^{son} = \alpha + \beta \cdot Y_i^{father} + \varepsilon_i.$$

In equation (1), Y denotes the natural logarithm of mean income. As commonly done by other studies, I also control for the son's and father's age and age squared to account for their different stages in the life cycle. The coefficient of interest β is the IGE, which indicates the extent to which father's permanent income level affects his son's. The higher the β, the more likely sons will inherit father's economic position and the lower the intergenerational mobility. In the extreme, when β equals zero, father's income has no bearing on the son's and there exists perfect income mobility. In contrast, if β is greater than or equal to one, not only economic status tends to be passed down to the next generation, but the income distribution fails to regress to the mean.

The regression investigates the net effect of father's income on the son's through any possible channels, obviating the need to include other control variables on the right-hand side without incurring the missing variable bias. But researchers have been trying different ways to handle the attenuation bias caused by the measurement error in father's permanent income. Single-year income is not a good proxy as it consists of both permanent income and transitory fluctuations (Solon, 1992; Solon, 2002; Mazumder, 2005). Three methods to mitigate the bias have been recommended. One is to take the average of the incomes across several years (usually 3 to 5 years) to get a more accurate measure of permanent income (Solon, 1992; Zimmerman, 1992; Bjorklund et al., 2012). However, Mazumder (2005) argues that the 5-year average income still suffers from large measurement error, as transitory fluctuations are probably persistent. Second, if longitudinal income data are not available, instrumental variable regressions can be used. Among the most well-known IVs are a father's education (Solon, 1992; Dearden et al., 1997) and a father's social or economic status (Zimmerman, 1992; Dearden et al., 1997). These two instruments have fewer transitory fluctuations than the current incomes. Using them as the IVs alleviates the downward bias. IV regression, however, is also problematic in that the IV may be invalid. Take the father's education level as an example: if a father's education is somehow positively correlated with his son's income after controlling for his own income, the IGE will be overestimated. The third method uses the two-sample two-stage least squares (TS2SLS) procedure (Bjorklund and Jantti, 1997; Dunn, 2007). Three steps are taken. First, the relationship between permanent income and personal characteristics is established using a complementary data set. Then, with the estimated relationship, father's lifetime income can be predicted using characteristics in the primary data set. Finally, the child's income is regressed on the predicted father's income.

The measurement error in the son's income does not in and of itself cause biases, but many researchers still choose to average their incomes across years to gain greater efficiency (Gustafsson, 1994; Couch and Dunn, 1997; Mazumder, 2005). Another issue regarding the regression is that the IGE tends to be biased downwards when the sons are at the beginning stage of their career (Solon, 2002; Haider and Solon, 2006). The correction of this problem is to only use sons within a specific range of age, typically from their late 20s to early 40s, or to average their earnings only in their latest years in the data such that the current income is close to the lifetime income (Gustafsson, 1994; Jantti and Osterbacka, 1999; Bjorklund and Jantti, 1997).

I follow the literature and use the average income across all available years as a measure of father's lifetime income. For the sons, I take the average of their incomes between 25 and 40 years old. I also run regressions using sons of all ages to take advantage of the larger sample size.

Additionally, I perform IV regressions for comparison. The instrumental variables I use are father's years of education and the average income within father's occupation. The former IV is the most commonly used in the literature. The rationale for the latter one is that fathers' incomes are highly correlated with their occupations, but the occupation itself does not directly affect their sons' income. As such, the exclusion condition is well satisfied. However, note that if father's occupation has predicting power of son's occupation, which affects son's income, the result will be biased upwards.

In this sense, the IV regressions provide an upper bound of β. The first-stage regression shown in Table 1.3 indicates that the IV is very strong.

1.3.2. Results

Table 1.4 presents all of the results. OLS using sons of all ages indicates that China's IGE is approximately 0.5. Restricting the sample to older sons reduces the estimate to about 0.4. As has been found in the literature, using IV regressions increases the estimates substantially, to between 0.59 and 0.80. Given that OLS tends to underestimate and IV regression tends to overestimate the IGE, the true value may lie between 0.5 and 0.6. In addition, since the majority of the fathers have income observations in 7 or 8 waves out of a total of 8 waves, the average income should be quite representative of their long-run incomes and the measurement error is supposed to be effectively wiped out.[5] As such, China's IGE is likely to lean toward the OLS end, and settle near 0.5. By comparing the results from OLS and IV regressions with the corresponding estimates in Table 1.1, I find that China's IGE is greater than those of most developed nations, especially the Scandinavian countries. The implied low mobility might be due to more nepotism and rent-seeking in recent years in China (Yuan and Lin, 2013).

There are big differences between China's urban economy and rural economy in many aspects. To explore whether it is true for income mobility, I split the sample into urban subsample and rural subsample according to father's residence and estimate the IGE for these two areas respectively. Since the number of observations (especially for

[5]On average, fathers report their incomes 6.65 times, with the survey waves spanning 17.71 years. About 60% of fathers have 7 or 8 income observations.

the urban areas) is fairly small, I use sons of all ages in the regressions to maintain a reasonable sample size. Table 1.5 reports the OLS results. It turns out that the urban households have a much higher IGE than the rural ones. It is not unexpected given that tens of millions of Chinese rural people, most of whom are young men (second generation), have migrated into cities to work in the manufacturing industry during the past two decades, which improved their earnings and weakened the link between father's and son's incomes. A further investigation into the data confirms this phenomenon: 84% of the urban working sons live at home while the number for rural working sons is only 62%.

1.4. The Distributional Pattern of Income Mobility

One interesting question concerning mobility is whether rich parents and poor parents have an equal effect on their children's future incomes. If not, the mean IGE misses a lot of information.

Previous papers have adopted five different approaches (transition matrix, linear spline regression, nonparametric regression, quantile regression and instrumental quantile regression) to examine the distributional pattern of mobility in other countries. In the following section, I apply all these methods to study this question in China. It is important to maintain a relatively large sample size because all of these methods allow mobility to vary across cohorts. Thus, I do not impose restrictions on son's ages. It may bias the estimates of mobility upwards, but to the extent that every cohort is similarly affected, it will not change its distribution pattern.

1.4.1. Transition Matrix

One of the most popular methods of addressing this question is with the use of a transition matrix (Jarvis and Jenkins, 1998; Corak and Heisz, 1999; Bratberg et al., 2005; Shi et al., 2010). Using my data, I construct the following quintile matrix:

It is notable that for the richest and poorest parents, their children are more likely to maintain the same income level while children with parents who fall in the middle groups enjoy much more income mobility. For example, about 43% ((5, 5) entry of the matrix) of the sons born to top rich fathers will end up being the richest as well. On the contrary, for the three middle-income cohorts, the probability of children having the same income status as their parents is well below 30%.

However, people are concerned that the way the matrix is constructed guarantees higher numbers at the end points and smaller numbers in the middle (Atkinson et al., 1983; Corak and Heisz, 1999).[6] This ceiling/floor problem is recognized in the early 1980s; unfortunately, researchers keep using the matrix without trying to address this flaw.

In this paper, to address this concern, I take two steps to modify the matrix. First, I divide parents and children into ten groups respectively and expand the matrix to a 10×10 one. The transition matrix becomes:

[6]Consider a father in the lowest quintile. If his son's income improves substantially, this father-son observation will wind up in a position of (1, 2), (1, 3), (1, 4) or (1, 5) entry in the matrix, depending on the extent of the enhancement. In contrast, if the son has a worsened income compared to other children, no matter how poor the situation is, he will still stay in (1,1) of the matrix. The situation is different for a middle-class family because they will not stay on the diagonal of the matrix whenever the son has a considerable change in his income, either richer or poorer. For example, a son with third-grade parents could move to (3, 4) or (3, 5) if he enjoys a higher earning, or switch to (3, 1) or (3, 2) if he makes less money. In other words, the poorest and the richest have only one direction of change while people in between can go either way.

It is conspicuous that the top decile (the (10, 10) entry) has an exceptionally large number, much larger than its bottom counterpart, indicating a lower mobility for the richest families than the poorest families. As a second step, I define a new concept called “relatively stable”. It means that the son's income either stays in the same decile as his father's income or moves to the neighboring deciles. For instance, the probability that a father in the bottom 10% stays “relatively stable” is the sum of the (1, 1) and (1, 2) entries whereas the chance of a “second to poorest” father (10%-20%) remaining “relatively stable” is the sum of the (2, 1), (2, 2) and (2, 3) entries of the matrix. As such, the ceiling/floor effect associated with the richest and poorest families is counteracted by only adding up two numbers as opposed to three. Table 1.8 and Figure 1.2 summarize these probabilities for all the ten subgroups.

The large numbers for the top 20% families imply that sons with rich fathers can easily inherit their favorable economic status. It should not be surprising as wealthy parents have more power and resources to transmit their income advantages. On the other hand, people in the left tail (lowest decile) of the income distribution enjoy relatively high mobility. This pattern comes as a consolation since it means that the moving up mechanism is not blocked for children from poor families. Given that the rural areas are typically poorer than the urban areas, this distributional pattern is also consistent with the previous finding that mobility in the rural areas is higher than in the urban areas.

1.4.2. Linear Spline Regression

The second method I use is linear spline regression. It allows the slope of the regression line to change at each pre-defined breakpoint (called knot), which can be designated arbitrarily. The regression yields a series of coefficients in each interval created by any two neighboring knots. By observing these coefficients, one is able to know how the coefficient of interest varies across the quantiles of the explanatory variable. Mathematically, suppose we separate the regressor into two pieces connected at z, and y is piecewise regressed on x, the regression equation would be:

$$y_i = \alpha + \beta \cdot x_i + \gamma \cdot (x_i - z) + \varepsilon_i. \tag{2}$$

In this paper, I define the knots as the 20th, 40th, 60th and 80th percentiles of father's log average income. I do not divide it into finer pieces because the coefficients will become overly volatile when the sections are small. The results are presented in Table 1.9 and Figure 1.3. Generally speaking, the IGE is the lowest for the bottom twenty percent fathers and rises to about 0.7 before declining slightly to 0.56 for the top twenty percent. This pattern also indicates greater mobility for the poor families than the wealthy families as is suggested by the transition matrix.

1.4.3. Nonparametric Regression

Nonparametric technique allows very flexible coefficients throughout. It does not make assumptions about the functional form of the relationship between the dependent and independent variables.

There are various specifications for nonparametric regressions. In this work, I adopt kernel weighted local polynomial smoothing, set the degree of the polynomial equal to 3 for the functional form, choose Epanechnikov kernel, set the bandwidth equal to 0.8 and abandon the outliers at the left tail as they are not likely to represent people's lifetime income.[7] The result is shown in Figure 1.4 where the blue line is the fitted line. Since the horizontal axis and vertical axis represent log mean father's income and log mean son's income respectively, the slope of the fitted line (which is also the derivative of log mean son's income with respect to log mean father's income), is the estimated IGE. It is obvious that the slope starts being zero or even negative in the first place. Then it turns positive and grows monotonically until around the midpoint. After that, the slope keeps roughly constant. It implies that the poor families are more mobile than the rich ones. This pattern further corroborates the findings in 4.1 and 4.2.

Summing up the results so far, the transition matrix, the linear spline regression, and the nonparametric regression all suggest considerably higher income mobility for the poor than for the rich. However, the implications for the middle-income families are not consistent across methods. The transition matrix shows that the IGE for the middle class is somewhat lower than for the poorest and much lower than for the richest group; the nonparametric regression indicates a higher IGE for the middle class than for the poorest, but there does not seem to be much difference between the rich and the middle class; finally, the spline regression shows that the IGE for the middle class is much higher than the poor and is even slightly higher than the richest group. The noise in the middle-income cohort could come from different econometric models and from the fact

[7] I have also used other nonparametric regressions such as locally weighted scatterplot smoothing (Lowess), and experimented with other bandwidths and kernels. The results are similar.

that the sample is not particularly large. The only conclusion that can be safely drawn from all three methods is that there is greater mobility for the poor than for the rich.

1.4.4. Quantile Regression

The remaining two methods to explore the distributional pattern of mobility are Quantile Regression and Instrumental Quantile Regression. Quantile regression has a few favorable properties compared with least squares. It does not require any distributional assumption of the error term and is robust to extreme values and outliers. This attribute is especially important for handling survey data since unusually large or small incomes are not rare. Another advantage is that it utilizes all the observations when computing the coefficients for each quantile without diminishing the sample size. This property is particularly useful for studying the distribution pattern of the IGE as people do not need to worry about the subsamples being too small.

The results of the quantile regression and IV quantile regressions are reported in Table 1.10 and Figure 1.5. A common feature these regressions share is that the richest sons happen to have low IGE and hence high mobility. The generally growing mobility across son's incomes (except for the lowest decile) is in accordance with Eide and Showalter (1999) who find a similar pattern in the USA.

1.4.5. Reconciliation of Different Results

Quantile regressions demonstrate that rich families end up with low intergenerational elasticity, which is at odds with what is implied by the transition matrix, the linear spline regression, and the nonparametric regression. To reconcile these results, first note that quantile regressions group people according to son's income whereas other methods are all based on father's income. They look at the problem from different perspectives.

Figure 1.6 provides a technical explanation of why decreasing IGE from quantile regressions and increasing IGE from a non-quantile regression can actually coexist. Suppose the observations are evenly distributed like a right triangle in the graph. According to the definition of quantile regression, the fitted line of a particular quantile in this case is simply a straight line that connects the series of points associated with a given quantile value for the dependent variable at each value of the independent variable. At the rightmost point of the triangle, every quantile value clusters there. Thus, all quantile regression lines must cross that point. It can be shown that the slope of the regression line decreases as the quantile gets higher (In the graph, the 80th percentile line must be flatter than the 20th percentile line. When the quantile reaches the upper bound 100th percentile, the slope is zero). By contrast, if the sample is divided conditional on the independent variable, and OLS is conducted in each subsample, the fitted line will start from being flat (in the left extreme, the slope is zero) and becomes steeper as the independent variable increases. Actually, Figure 1.4 depicts the distribution of the observations, which does look somewhat like the triangle described above.

A potential economic reason for the high income mobility associated with the wealthiest sons is that they may come from a broad range of economic backgrounds, such that the father's income does not have large explanatory power in regards to son's income. To test whether this is the case, I calculate the standard deviation of log father's average income for sons of different income levels. Note that taking the natural log takes care of heteroscedasticity in the father's income. Table 1.11 shows that fathers of the richest sons (top 2 deciles) indeed have a broader range of incomes than other fathers except for the poorest cohort. This should come as no surprise because China’s opening-up and reform policy, the adoption of the market economy and the compulsory education have given Chinese young men a great number of opportunities to build wealth, even if they do not come from rich families.

Finally, integrating the results from all these econometric techniques, one may conclude that: on one hand, rich parents can easily pass wealth to the next generation (according to the non-quantile regression methods); on the other hand, young men have various ways other than being born into an affluent family to become wealthy (according to the quantile regressions).

1.4.6. Comparing the Methods

There is no answer as to which method should be used when studying the distributional pattern of the IGE. However, it helps to keep in mind the shortcomings of each method.

Transition matrix suffers from two drawbacks. One is the ceiling/floor effect as is mentioned above. The second is missing information on ages of both fathers and sons, which can be problematic. Solon (2002) and Haider and Solon (2006) show that the IGE tends to be underestimated if sons are very young while Grawe (2004) illustrates that estimates of IGE are negatively correlated to father's age.

Linear spline regression allows people to choose the knots (specifying intervals) arbitrarily. However, the result is very sensitive to the choice of the knots. Poorly selected knots can result in exceptionally large coefficients at the cost of unusually small or even negative coefficients for the neighboring intervals.

Nonparametric regression also fails to directly control for ages. Besides, the best-fitting line may be affected by the outliers in the tails of the distribution. Finally, setting parameters may be challenging.

Quantile regression approaches a problem from the angle of the dependent variable, making it seemingly contradictory to other methods sometimes.

In sum, every method has its own drawbacks. Even with the same data, different methods can generate very different results. In the absence of theories, one should be very careful about putting too much weight on any single result; Robustness checks are necessary. It may be worthwhile to adopt multiple methods to get a range of coefficients before making conclusions.

1.5. Concluding Remarks

This paper explores intergenerational income mobility and its distributional pattern in China using a longitudinal sample. I find that China's IGE is likely to be between 0.5 and 0.6, which hints that China has less mobility than most of the developed countries. This conclusion is in line with the conjecture that developing countries provide people with less equal opportunity and are thus less mobile. Additionally, I adopt a variety of strategies to investigate the distribution of income mobility across income levels. It turns out that the modified transition matrix, the linear spline regression, and the nonparametric regression end up telling a fairly consistent story that the poor households in China are much more mobile than the rich, which implies opportunities for the poor children to escape poverty. However, quantile regression uncovers a different pattern that rich sons are actually a highly mobile cohort. Combining all the information I conclude that in China, while wealthy parents can easily make their children wealthy, there are plenty of ways for a child to become rich. Being born into an affluent family is not the only one.

Table 1.1: Selected Literature on the Estimates of the IGE in Different Countries

Country	**Study**	**Methodology**	**Estimated IGE**
Developed Countries			
Canada	Corak and Heisz (1999)	OLS	About 0.2
Finland	Osterbacka (2001)	OLS	0.13
France	Lefranc and Trannoy (2005)	TSIV	About 0.4
Germany	Couch and Dunn (1997)	OLS	0.12 if sons are 18+ 0.30 if sons are 25+
Great Britain	Dearden et al. (1997)	OLS, IV	0.24 using OLS, 0.39-0.44 using OLS with predicted wages, 0.56-0.59 using IV
Italy	Piraino (2007)	TS2SLS	0.55 or 0.44 depending on the definition of income
Norway	Bratburg et al. (2005)	OLS	0.129 for cohorts born in 1950; 0.155 for cohorts born in 1960
Spain	Pascual (2009)	OLS and IV	0.32 using OLS; 0.41 using IV
Sweden	Bjorklund and Jantti (1997)	TS2SLS	0.28
Sweden	Osterberg (2000)	OLS	0.13
USA	Solon (1992)	OLS and IV	0.41 using OLS; 0.53 using IV and single-year income in 1967
USA	Mazumder (2005)	Tobit	Around 0.6
Developing Countries			
Brazil	Andrade et al. (2003)	TSIV	0.60
Brazil	Dunn (2007)	TSIV, OLS, IV	0.69 using TSIV;0.53 using OLS;0.69 using IV
China (urban)	Gong et al. (2012)	TS2SLS	0.56 for all children; 0.63 for children aged 30-42
China	Fan et al. (2013)	OLS	0.32 for cohorts born between 1949 1970, 0.44 for cohort after 1970
Ecuador	Grawe (2004)	TSIV	1.13
Malaysia	Grawe (2004)	TSIV	0.54
Nepal	Grawe (2004)	TSIV	0.32
Pakistan	Grawe (2004)	TSIV	0.24
Peru	Grawe (2004)	TSIV	0.67

Notes: TSIV represents two stage IV regression.

Table 1.2: Summary Statistics

Total Sample		Mean	Standard Deviation
Number of Father-Son Pairs		1407	
Son	Average Annual Income (2009 RMB)	7029	13593
	Age in 2000	24.1	6.7
	Years of Education	9.28	2.5
	Living at Home (1=yes)	0.66	0.48
Father	Average Annual Income (2009 RMB)	5282	5514
	Age in 2000	52.8	8.6
	Years of Education	5.91	3.3
Son's Age is Restricted to 25-40 Years old		**Mean**	**Standard Deviation**
Number of Father-Son Pairs		442	
Son	Average Annual Income (2009 RMB)	8402	11098
	Age in 2000	28.1	5.5
	Years of Education	9.68	2.8
	Living at Home (1=yes)	0.78	0.41
Father	Average Annual Income (2009 RMB)	4826	3782
	Age in 2000	56.6	7.6
	Years of Education	5.41	3.3

Source: China Health and Nutrition Survey (1989-2009).

Table 1.3: First-Stage Regressions for IV=Occupational Income

Dependant Variable: Ln (Father's Mean Income)		
	All Ages	25-40 Years Old
Ln (Occupational Income)	0.295***	0.351***
	(0.03)	(0.05)
R2	0.2324	0.3155
F-test Statistic	84.84	40.19

Note: ***denotes significance at the 1% level.

Table 1.4: Estimates of China's Overall IGE

Covariates	Ages of Sons					
	All ages			25-40 years old		
	OLS	IV: Father's Education	IV: Father's Occupation	OLS	IV: Father's Education	IV: Occupation
Ln (Father's mean income)	0.498***	0.800***	0.717***	0.382***	0.594*	0.682***
	(0.051)	(0.227)	(0.188)	(0.071)	(0.327)	(0.230)
Age_{son}	-0.001	-0.012	-0.009	-0.214***	-0.186**	-0.174**
	(0.030)	(0.032)	(0.031)	(0.073)	(0.076)	(0.083)
Age^2_{son}	0.000	0.001	0.001	0.003**	0.003**	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age_{father}	-0.073	-0.062	-0.065	0.097	0.096	0.095
	(0.045)	(0.046)	(0.045)	(0.067)	(0.068)	(0.070)
Age^2_{father}	0.001	0.001	0.001	-0.001	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
F-statistic	29.32	11.01	11.65	20.87	13.51	15.67
R2	0.108	0.081	0.094	0.237	0.217	0.196
Observations	1407	1407	1407	442	442	442

Notes:

[a]The sample is from the China Health and Nutrition Survey (CHNS) 1989-2009.

[b]Dependant variable: Ln (son's mean income).

[c]Robust standard errors adjusted for within-household correlation are in brackets. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Table 1.5: Difference between the Urban and Rural Areas

Covariates	Rural	Urban
Ln (Father's mean income)	0.407***	0.694***
	(0.059)	(0.099)
Age_{son}	-0.016	0.062
	(0.040)	(0.059)
Age^2_{son}	0.001	-0.001
	(0.001)	(0.001)
Age_{father}	-0.053	-0.188*
	(0.049)	(0.107)
Age^2_{father}	0.004	0.002*
	(0.004)	(0.001)
F-statistic	17.19	12.17
R2	0.075	0.262
Observations	1189	218

Notes:
[a]OLS is employed. The dependant variable is the natural logarithm of son's mean income.
[b]Sons of all ages are used.
[c]Rural/urban is defined by parents' residence.
[d]Robust standard errors adjusted for within-household correlation are in brackets. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.

Table 1.6: 5×5 Father-Son Income Transition Matrix

Father's Income	Son's Income: Bottom	Second	Third	Fourth	Top
Bottom	0.326	0.259	0.216	0.135	0.064
Second	0.212	0.283	0.226	0.155	0.124
Third	0.196	0.206	0.210	0.221	0.167
Fourth	0.159	0.155	0.223	0.244	0.219
Top	0.107	0.096	0.125	0.246	0.427

Table 1.7: 10×10 Father-Son Income Transition Matrix

		Son's Income									
		Bottom	2nd	3rd	4th	5th	6th	7th	8th	9th	Top
Father's Income	Bottom	0.170	0.170	0.128	0.099	0.170	0.064	0.050	0.078	0.035	0.035
	2nd	0.128	0.184	0.156	0.135	0.071	0.128	0.085	0.057	0.043	0.014
	3rd	0.098	0.098	0.140	0.147	0.112	0.112	0.091	0.077	0.070	0.056
	4th	0.107	0.121	0.136	0.143	0.064	0.164	0.079	0.064	0.064	0.057
	5th	0.136	0.114	0.143	0.100	0.100	0.064	0.114	0.071	0.050	0.107
	6th	0.071	0.071	0.078	0.092	0.142	0.113	0.135	0.121	0.092	0.085
	7th	0.077	0.092	0.085	0.056	0.113	0.092	0.134	0.134	0.162	0.099
	8th	0.085	0.064	0.071	0.099	0.106	0.113	0.106	0.085	0.121	0.149
	9th	0.057	0.043	0.057	0.071	0.093	0.064	0.100	0.171	0.193	0.150
	Top	0.071	0.043	0.007	0.057	0.050	0.043	0.106	0.113	0.234	0.277

Table 1.8: Probabilities of Being "Relatively Stable"

0%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	90%-100%
0.340	0.468	0.385	0.343	0.264	0.390	0.430	0.312	0.514	0.511

Table 1.9: Linear Spline Regression Results

Quantiles	0-20%	20-40%	40-60%	60-80%	80-100%
Estimates of the IGE	0.277	0.443	0.720	0.700	0.559
F-statistic	19.46				
R2	0.112				
Observations	1406				

Table 1.10: Quantile Regression Results

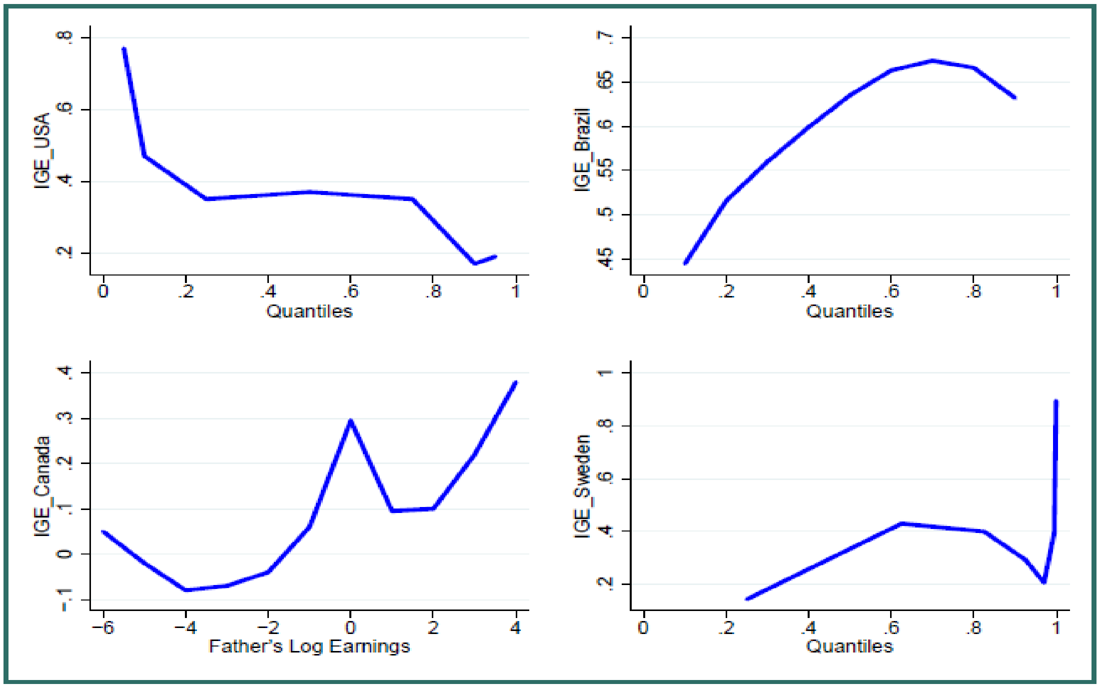
		IV Quantile Regressions	
Quantile	Regular Quantile regression	IV: Father's Years of Education	IV: Mean Occupational Income
0.1	0.390***	0.387***	0.839***
0.2	0.574***	0.870***	0.826***
0.3	0.527***	0.759***	0.970***
0.4	0.569***	0.886***	0.863***
0.5	0.536***	0.973***	0.924***
0.6	0.485***	0.782***	0.670***
0.7	0.443***	0.760***	0.546***
0.8	0.437***	0.589***	0.587***
0.9	0.415***	0.665***	0.670***

Notes: ***denotes significance at the 1% level.

Table 1.11: The Variation of Father's Income for Different Groups of Sons

Percentile of Son's Income	0%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	90%-100%
Std Dev of Log Father's Income	0.75	0.64	0.54	0.63	0.69	0.60	0.64	0.63	0.74	0.69

Figure 1.1: Distributional Patterns of the IGE in Other Countries



Notes: For USA and Brazil, quantile regressions are applied. So the horizontal axes represent quantiles of the children's incomes. The remaining two use non-quantile regression techniques. The horizontal axis in the graph for Sweden represents quantiles of the parents while in the Canada case, it is father's log earnings. It is noteworthy that these results cannot be directly compared due to different econometric methods employed. As is shown in Section 4, quantile and non-quantile regressions could generate opposite results even if the same data set is used.

Figure 1.2: Results of the Modified Transition Matrix

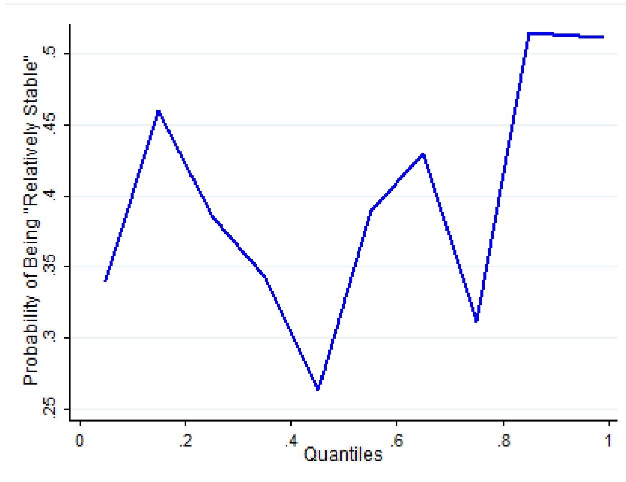


Figure 1.3: Results of the Linear Spline Regression

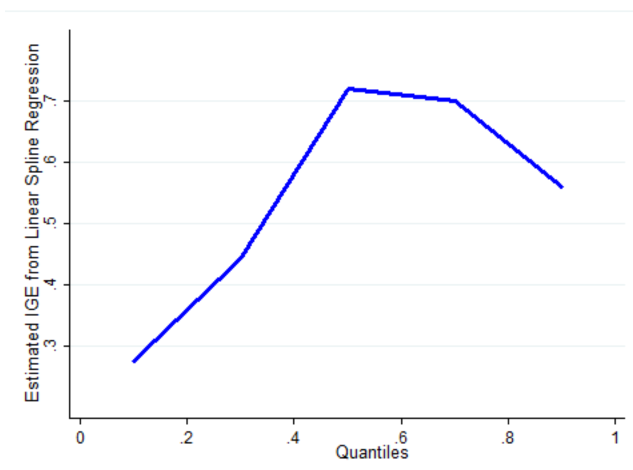


Figure 1.4: Results of the Nonparametric Regression

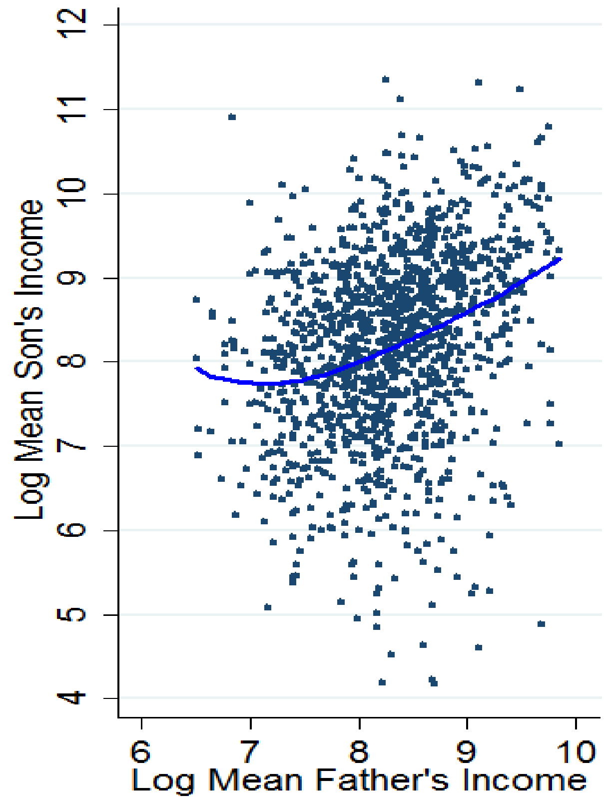


Figure 1.5: Results of the Quantile Regressions

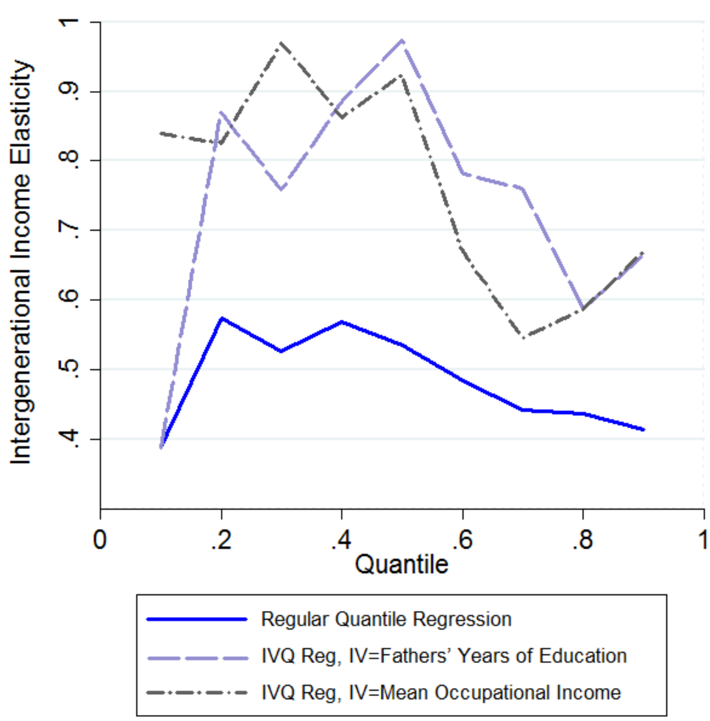
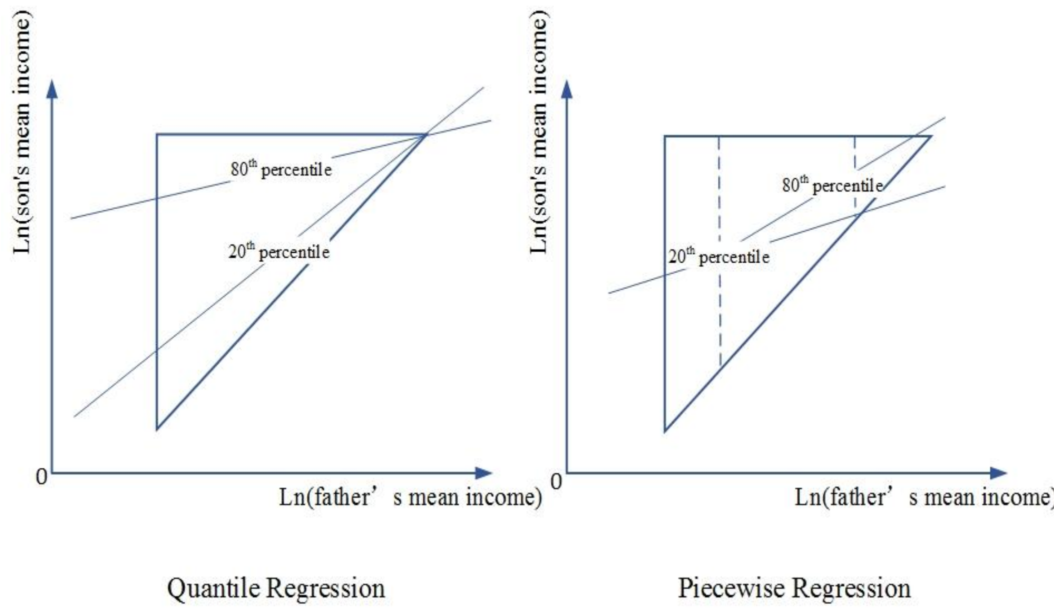


Figure 1.6: Analysis of the Difference between Quantile and Piecewise Regression



Chapter 2: Marriage Market Matching and Conspicuous Consumption in China

2.1. Introduction

A preference for sons, coupled with the one child-policy, has combined to generate a relative shortage of females in China. Estimates from China's Population Census suggests that for cohorts born over the period 1970-2000, males have grown in share from 51% to 57% of the total population (Qian, 2008, p1251). Estimates of the share of "missing women" in the country now exceed 40 million women (Bulte et al., 2011). Existing research has shown that these skewed sex ratios are producing widespread social and economic upheaval, including higher saving rates, ballooning housing prices (Wei and Zhang, 2011; Wei et al., 2012), higher rates of bachelorhood (Guilmoto, 2011), and reduced overall welfare (Bhaskar, 2011). Recent research also suggests that spending on status goods appears to be growing in magnitude, even in poorer parts of rural China (Brown et al., 2011; Chen et al., 2012).

In this paper, we investigate the extent to which skewed sex ratios influence expenditure decisions in a broad sample of roughly 24,000 automobile transactions across China over the period 2009-2011. Using a differences in differences in differences approach, we show that unmarried male consumers who reside in an area with an unfavorable sex ratio purchase more expensive vehicles than their married peers. We also identify specific luxury vehicle models and confirm that the cars purchased by these consumers are more likely to be classified as higher end models.

The pressure generated by skewed sex-ratios on individuals to consume conspicuously may vary along a number of dimensions. We are able to explore two sources of heterogeneity in our sample. First, we show that the estimated relationship is both larger in magnitude and more precisely estimated among the quartile of individuals living in areas with the most highly skewed sex ratios. Second, we find that the poorest quartile of unmarried males in our sample exhibit the largest changes in expenditure.8 This result is consistent with a situation in which individuals of varying income (and potentially of varying reference groups) use different sets of commodities to signal wealth, as documented by Chai and Kaus (2012) for South African consumers. Finally, we investigate whether this behavior generates negative externalities in the form of lower average fuel economy and higher average vehicle weight, a factor that can significantly increase mortality in traffic accidents (Anderson and Auffhammer, 2014). Our results suggest that consumption signaling skews the pool of purchased automobiles to lower gas mileage vehicles, but has little impact on overall vehicle weight.

The use of consumption expenditure to signal social status has attracted a great deal of attention in the economics literature. Recent work has focused on empirically identifying consumption visibility and exploring the determinants and motivations for signaling (Charles et al., 2009; Heffitz, 2011). These efforts routinely classify automobiles as among the most conspicuous of purchases and a number of papers focus specifically on vehicle purchases. For instance, Grinblatt et al. (2008) show that Finnish consumers are directly influenced by the automobile purchases of their nearest neighbors, particularly for recent purchases.

[8] Our sample of car buyers is more affluent than the population at large. We explore the sample at length in Section 2.

Positional spending has also been studied in developing countries, with research highlighting the fact that the allocation of expenditure for this purpose has the potential to act as a poverty trap in this setting (Banerjee and Duflo, 2007; Brown et al., 2011; Moav and Neeman, 2010, 2012; Case et al., 2013; Kaus, 2013). These effects appear to be amplified by marriage market conditions. For instance, Chen et al., (2012) show that relative to families with daughters, poor Chinese households with sons undertake higher levels of social gift-giving associated with maintaining guanxi, networks of influence in Chinese society.

Our analysis sample confers several unique advantages in this context. First, we examine expenditure on automobiles, a ubiquitous and highly visible commodity, with a range of purchase options, and for which signaling is often considered a major consumption motivation. Second, our study spans consumers across all of China, both rural and urban. Finally, because the data comes from a financial lender, our data contains information about the products purchased as well as detailed records on the consumers themselves.

At the same time, because we are working with a very specific set of consumers and consumption decisions, the generalizability of the results we obtain may be limited. In particular, we are identifying a relationship between positional spending and marriage market competition within a subsample of Chinese car consumers, who are affluent enough to buy a car but both need to borrow to make their purchase and engage with this specific lender. Estimates from the 2011 Chinese Household Finance Survey suggest that roughly 27% of new car purchasers in China rely on credit. To the extent that richer households are less likely to need credit, our estimates are thus simply best interpreted as the impact of the sex ratio among the population of purchasers who choose to employ credit.

Rates of private car ownership in China are low but are increasing rapidly. While in 1985 the rate of ownership was only 0.27 for every 1000 people, this rate has risen to 55 per 1000 in 2011 (Feng and Li, 2013). Rising affluence combined with growing automobile ownership suggests that the scope for status competition through vehicles and the effects that we observe may intensify in the future. The extent to which our findings apply more broadly depends on the extent to which this group is selected and to which motivations for consumption signaling could vary across other commodities, over time, or by income level.

Our paper is organized as follows. Section II describes the construction of our dataset and presents summary statistics on the sample investigated. Section III undertakes the differences in differences analysis and discusses the implications of marriage market competition for consumption behavior in this setting. Section IV explores the potential for heterogeneity in the estimated relationship and examines several potential implications of consumption signaling behavior specific to the automobile industry. Section V concludes.

2.2. Data

The data is constructed from three principal sources. Information on automobile transactions and car purchasers themselves is provided by a large Chinese financial institution from 2009 through 2011. The source provides records on 24,134 individual loans, and includes borrowers from all provinces of mainland China except Tibet. A key advantage of this data is the level of detail contained on the vehicles purchased and on individual borrowers including information on marital status, age, gender, education, and earnings as well as on the geographic location of purchase. We also derive information on prefecture-level outcomes, which are obtained from both provincial statistical yearbooks or from the National Bureau of Statistics (CEInet, 2015).

Panel A of Table 2.1 presents summary statistics on car buyer characteristics. One noticeable feature is that borrowers are comparatively rich, with average annual earnings of nearly 100,000 RMB, which is over $16,000. In contrast, average income per capita is roughly 35,100 RMB according to the National Bureau of Statistics of China, which highlights that the sample of car purchasers is highly selected.[9] Most purchasers are male, middle-aged, married and have received at least a college education. Panel B provides detail on the cars purchased in the sample.[10] As can be seen from the table, over half of the cars are foreign brands and approximately one-fourth are luxury vehicles. The average miles-per-gallon (MPG) is city/highway combined and is slightly below 30.

Panel C reports prefecture-level characteristics. The dataset contains transactions occurring in some 292 of the 334 total prefectures in China, while we have prefecture-level economic statistics for roughly 254 regions. Across regions with data, average income is 45,324 RMB. The mean GDP per capita across prefectures in our sample is higher than China's GDP per capita in 2011 (35,100 RMB), a difference which reflects

[9]Comparable estimates from the 2011 Chinese Household Finance Survey suggest that new car purchasers earn on average 65,000 RMB, which is closer to that found in our sample. The true difference may be even smaller as our data is not survey data, and it is well known that survey estimates of income, particularly in developing countries often understate mean income (Deaton, 2005).

[10]Expenditures have been deflated to real prices using a base month of September 2010 and the official consumer price index.

the fact that more automobile purchases are undertaken by inhabitants of wealthier regions and that this value is not population weighted.[11] The overall sex ratio, obtained from Provincial Statistical Year books, based primarily on Hukou derived estimates, is roughly 1.05 males per female across all prefectures, although this value is more skewed for younger cohorts. We explore the robustness of our analysis to alternative measures of the sex ratio, such as those based on Census data, in Section 3. Furthermore, this value masks substantial geographic heterogeneity; Figure 2.1 depicts the sex ratio by prefecture (measured as the ratio of males per 100 females). Although some of the most severely imbalanced gender ratios exist in rural areas, the issue is not limited to such areas. Overall, the sex ratio ranges from a low of 90 to a high of over 132 men per 100 women.

2.3. Analysis

Estimating the relationship between the level of competition in the marriage market (here measured by the sex ratio) and positional spending (captured by the purchases of more luxurious automobiles) is complicated by a number of factors. For example, a primary concern is that the presence of a more skewed sex ratio may be correlated with other omitted factors such as variation in terms of average incomes, education levels, or even male/female gender roles in a region.

We attempt to mitigate this and related sources of endogeneity through the use of a triple differences approach. Specifically, we compare male and female car purchasers,

[11]The financial institution may also be more likely to make loans to individuals in richer prefectures. In order to maintain a representative sample of Chinese consumers, we exclude those with annual incomes in excess of 250,000 RMB.

who are and are not married, across low and high sex ratio regions. Our identification of the effects of high sex ratio on Chinese men's car-purchasing preference thus relies on the interaction between gender and marital status. In Chinese culture, the family of a groom is traditionally responsible for purchasing a house at the time of marriage, suggesting that males may also have a stronger incentive than females at this stage of their life to signal wealth through a range of visible expenditures (Wei *et al.*, 2012). This pressure has intensified with increasing competition for brides and the incentive should act on the unmarried rather than on married men (Wei and Zhang, 2011).

For illustrative purposes, we begin our analysis with an example triple difference mean estimate presented in Table 2.2. For simplicity we calculate means using the top quartile and bottom quartiles of sex ratio.[12] In the more imbalanced areas, unmarried men, on average, purchase cars that are 7025 RMB (column (5)) more expensive than those unmarried women purchase. By contrast, for married individuals, men's automobiles are only 811 RMB (column (6)) more expensive than married women's purchases. The difference in difference is 6214 RMB as is shown in the top of column (7).

Next we carry out the same analysis for the quartile of less skewed sex ratio provinces. We find that the mean difference-in-difference here is small but actually negative (-996 RMB as shown in the middle of column 7), implying that there is no obvious car-related conspicuous consumption in this group. Finally, our triple difference estimate is simply the difference between these two double differences: 7210 RMB, (or around $1200) which is an economically important effect.

[12]The mean sex ratios in these two groups are 100.6 and 110.5 respectively.

In what follows, we run regressions that implement a regression version of this basic triple-difference approach. Equation (1) presents this triple difference estimation strategy, with all component interactions included:

$$\begin{aligned} \text{Carprice} &= \alpha + \beta_1 \text{male} + \beta_2 \text{unmarried} + \beta_3 \text{sexratio} \\ &+ \beta_4 \text{male} * \text{unmarried} + \beta_5 \text{male} * \text{sexratio} + \beta_6 \text{sexratio} * \text{unmarried} \\ &+ \beta_7 \text{male} * \text{unmarried} * \text{sexratio} + \theta \text{X} + \delta Z + \eta + \varepsilon \end{aligned} \tag{1}$$

where X is a vector of borrower characteristics and Z is a vector of prefecture and province level controls. In theory, the coefficient on the triple interaction, β_7, should isolate the specific portion of consumption expenditure undertaken by the group with the highest motivation to signal status, unmarried males facing high levels of competition. In some specifications, we also include a set of province fixed effects, denoted by η.

The results of estimating equation (1) are presented in Table 2.3 with each column (1)-(5) adding additional controls. Column (1) estimates the triple difference with no controls. Column (2) includes controls for borrower characteristics including age, age squared, earnings and educational attainment, while column (3) additionally includes the prefecture characteristics listed in Table 2.1. Finally, column (4) includes province fixed effects, while column (5) incorporates a control for housing prices (for regions of the China with available data) because houses have also been specifically shown to be a positional good in the Chinese case (Wei and Zhang, 2012). We include controls for local housing prices only in some specifications because they dramatically limit the sample.

In all specifications, the coefficient estimate of interest on the triple interaction term is positive and statistically significant at the .10 level or better. The estimated impacts are economically meaningful as well. To put the magnitude of the coefficients in perspective, the estimated coefficient on the triple interaction in Column (5) implies that if one were to go from a prefecture in which there is parity in the sex ratio to a prefecture with a sex ratio that is 1 standard deviation above the mean in favor of males, then the typical car purchased by an unmarried male would be 4,285 RMB more expensive. This change represents a 3.3% increase in the mean purchase price paid by this group of consumers.

An additional concern is that several large Chinese cities such as Beijing, Guangzhou, and Shanghai have also enacted vehicle ownership restrictions such as auctions and lotteries in an effort to curb congestion and pollution (Feng and Li, 2013). Restrictions may alter the pool of individuals eligible to purchase cars and may limit the scope for status competition through consumption signaling in these regions. As a check, Table 2.4 demonstrates that the results we obtain are robust to the exclusion of these areas.

Higher purchase prices could reflect numerous car characteristics, not all of which may equally convey status. An advantage of studying automobiles is that vehicles are already classified as standard or as luxury models both across and within producers. As an alternative approach, we consider the likelihood an individual purchases a car which is classified as luxury. To do this we estimate an equation as in (1), using a probit specification, with an outcome variable which is an indicator taking the value of one for luxury automobiles. The results of this exercise are presented in Column (6) of Table

2.3 and are reported as the marginal effect. Using the same comparison as in Column (5), the estimates suggest that unmarried male consumers in prefectures one standard deviation above the mean would be 4.8 percentage points more likely to purchase a luxury car than those in a balanced prefecture.

While the OLS results are instructive, the use of income as a regressor in the model makes it possible that we may have issues with heteroskedasticity.[13] In fact both a general test (the White test) and a specific test focusing on income (the Goldfeld-Quandt test) decisively reject the null of homoscedasticity at the 0.01 level (results not shown). As a result, we re-estimate our model in Table 2.5 using Feasible Generalized Least Squares (FGLS) where the error variance is proportional to income raised to an unknown exponent that we estimate in a first stage regression using the OLS residuals. As can be seen, our coefficient of interest, the interaction of unmarried, male, and the sex ratio is fairly constant in size across the two approaches, but estimated much more precisely using FGLS. The estimates now exhibit significance at the .05 level or better in all five specifications. We thus use the FGLS model approach for the subsequent analyses in Section IV.

As a final robustness check, we examine the sensitivity of our estimation results to alternative measures of the sex ratio. While our primary measure has the distinct advantage of having consistent data for a very large number of prefectures, alternative measures exist which can be used to disaggregate sex ratios by age group instead of those for the entire local populace. For instance, it is possible to calculate age-specific sex-ratios for many prefectures using China's National Census from 2010. This can also

[13]Heteroskedasticity related to scale is a classic case in the literature, and if the scale variable is also included as a regressor in the model, the efficiency gains from using GLS can be substantial. See Baum (2006, pp. 144-147).

be done at an aggregated level across all provinces. Because marriage market pressures should only be exerted by the presence of those of marriageable age, it is worthwhile to examine the local sex ratio of these groups specifically.

We reproduce our results using these alternative measures of local sex-ratios in Table 2.6. Columns (1) and (4) reproduce our original OLS and FGLS results for comparison. We present results for the age range 20-49 because most initial marriages fall within this range. We elect for this range because the legal age of marriage for women in China is 20.[14] Results using data we extracted from the 2010 Chinese National Census are presented for the OLS and FGLS regressions in (2) and (5).[15,16] In both cases, the magnitude of the effect is slightly smaller, but the sign and significance of the estimated impacts remain consistent. We also reproduce the analysis using the overall provincial age-specific sex ratios obtained from statistical year book data. These are presented in columns (3) and (6). This alternative approach yields slightly larger estimated impacts, but is again consistent with the original results.

2.4. Extensions

[14]Reassuringly, the results are also not overly sensitive to this choice. In addition to those presented in the table, we additionally examined multiple windows of age specific sex ratios, ranging from more narrow (20-39) to broader (15-64), and this exercise generally produces results consistent with those in the text -- although the magnitude and significance of the estimate coefficient varies slightly from specification to specification.

[15]This is perhaps not surprising as correlations between our population level measure and age-specific sex ratios are quite high (0.84 for the 20-49 year old sample for example).

[16]This data source is not available for as many prefectures as could be obtained from the provincial statistical yearbooks, particularly for less populated areas, so we apply the disaggregated age-specific sex ratios of the overarching province when this is unavailable.

2.4.1. Heterogeneous Effects

Results presented in Tables 2.3 through 2.5 suggest a highly significant conspicuous consumption effect, where every one point increase in the sex ratio raises car spending by roughly 450 RMB for unmarried males relative to all other types of individuals. In this subsection we consider two possible types of heterogeneity in this relationship. The first is heterogeneity across income levels. This is especially important in the context of conspicuous consumption where the positional goods used for signaling may vary for different reference groups and income levels (Chai and Kaus, 2012). For example, in our context, it is quite possible that, at very high levels of income, unmarried males use ownership of houses or land to signal their worthiness to potential partners (as demonstrated by Wei and Zhang, 2011) while those with lower income levels compete by signaling with relatively lower cost commodities such as automobiles.

To investigate, we split our data in quartiles of income and estimate a separate triple difference FGLS regression within each subgroup. These results are reported in Table 2.7 and strongly suggest that the use of cars as a signaling mechanism is concentrated primarily in the lower income quartile in our sample. Specifically, we document a highly significant average effect which is roughly 50% to 100% larger than our average effect estimated for the full sample. Estimates for the second income quartile are small, negative and not significantly different from zero. Those for the third and fourth income quartiles also show effects larger than our full sample estimates, but they are generally not precisely estimated, only occasionally reaching significance at the 0.10 level.

A second possibility we investigate is that pressure for consumption signaling may vary in a non-linear manner across regions as a function of the relative shortage of women. For example, it may be that a one unit change in the sex ratio in highly skewed regions may not have the same effect on the consumption pattern of unmarried males in only weakly skewed regions. It could be the case that at relatively balanced sex ratios, the incentive to compete by conspicuous consumption may be attenuated relative to the incentive at higher ratios.

We investigate this case in Table 2.8 which parcels the sample in quartiles of the sex ratio distribution. It is apparent that robust and precise estimates only appear for the most skewed sex ratio regions. In these regions sex ratios are highly skewed, ranging from 107 to 132 males per female, so it is unsurprising that this would be the subset of prefectures where pressure for consumption signaling is greatest.

Our sample is not large enough to split the data by both sex ratio and income quartiles at once and still precisely estimate the model. At the same time, the results above provide suggestive evidence that it is the relatively poorest males in the regions with the most imbalanced sex ratios that heavily use car purchases as conspicuous consumption to signal to potential marriage partners.

2.4.2. Externalities Associated with Conspicuous Consumption

A number of studies have established negative impacts of consumption signaling. This includes diversion of expenditures away from commodities thought to have positive externalities such as education, with particularly detrimental consequences for lower income households (Banerjee and Duflo, 2007; Brown *et al.*, 2011; Charles *et al.*, 2009; Moav and Neeman, 2010, 2012; Case *et al.*, 2013; Kaus, 2013). Other authors have argued that collective action may make conspicuous consumption an unproductive necessity. In other words, to the extent that everyone within a reference group undertakes some expenditure specifically to signal status, the ultimate result can be that individuals engage in a rat race with no net change in local distribution of status (Hopkins & Kornienko, 2004).[17] We explore a further potential impact of this behavior – that the expenditure changes may directly generate negative externalities.

Automobiles are a commonly studied commodity thought to generate negative externalities due to their impact on air pollution, congestion, and use of natural resources. Some of these can be quantified and measured in a standardized form. For our sample, we compile data on average mileage (in MPG) and weight (in lbs) of all vehicles in our data. MPG should be informative of the impact cars have on both resource use and on pollution, as well as the vehicle's lifetime usage cost. Weight can generate negative externalities, both through its impact on average MPG and through its impact on the severity of car crashes (Anderson and Auffhammer, 2014).

To get a sense for how large the impact of these externalities could be, we focus on prefectures in the highest quartile of skewed sex ratio and estimate the FGLS regression as in Table 2.8.[18] As can be seen from the Table 2.9, cars purchased by unmarried males in the most skewed regions exhibit both lower fuel economy and higher weight. The MPG effect is rather sizeable, with a one unit increase in the sex ratio among the most

[17] In such a world, there exists a pareto-improvement in which one could reallocate everyone's expenditure away from positional commodities in favor of others which may be desirable on other grounds, but doing so unilaterally is not individually rational, as failing to signal alone would net a relative fall in status.

[18] Estimates obtained from regressions utilizing the full sample are roughly one fourth to one half of those for this group in magnitude and are generally less precisely estimated.

skewed sex ratio regions being associated with a reduction in fuel efficiency in the range of 0.30 to 0.47 MPG. This suggests that consumption signaling in this setting may indeed be exacerbating this class of existing negative externalities associated with automobiles. Interestingly, the impact of vehicle weight is not economically meaningful in size with these vehicles being only a few pounds heavier per unit change in the sex ratio.[19] As can be seen in Panel B of Table 2.1, the standard deviation of weight is pretty low as well, suggesting that cars purchased in China are relatively homogenous in size.

2.5. Conclusion

Using a novel dataset on borrowers in China, we have shown that the increasingly skewed sex ratios in many parts of the country are creating incentives for unmarried men to significantly alter their automobile consumption habits in ways that appear competitive in nature. The effects that we observe are strongest in the quartile of regions with the most unbalanced sex ratios, suggesting that positional spending may yet worsen if the sex ratio continues to deteriorate.

We further demonstrate that the largest expenditure reallocations occur among the poorest quartile of borrowers in our sample. This evidence supports the hypothesis that the range of commodities used to jockey for social status varies across individuals in different portions of the income distribution (or among those facing different peer groups). Beyond the direct social inefficiency that status competition represents, we demonstrate that this behavior also leads to consumption of vehicles with lower vehicle fuel efficiency. Thus, to the extent that China’s one child policy has generated

[19] This suggests that it is possible to signal status through vehicle quality without requiring larger cars to do so.

incentives which further skewed sex ratios, it may also have exacerbated social status competition. In many areas of the country, unmarried men now appear to compete in a zero-sum game of consumption signaling, one which is capable of generating large negative externalities. Given that there is little reason to suspect that growth in the rate of car ownership in China will slow, conspicuous consumption may become even more important and generate greater aggregate distortions as time passes.

Table 2.1: Summary Statistics

Panel A: Borrower Characteristics	Obs	Mean	Std. Dev.
Age	24,133	35.62	7.91
Male	24,133	0.75	0.43
Earnings	24,133	98,027	51,608
Married	24,133	0.84	0.37
Educational Attainment			
Illiterate	24,133	0.00	0.02
Elementary	24,133	0.10	0.30
High School	24,133	0.30	0.46
College	24,133	0.59	0.49
Graduate	24,133	0.01	0.11
Panel B: Car Characteristics			
Purchase Price	24,133	131,071	46,724
Foreign	24,133	0.56	0.50
Luxury Make	24,133	0.24	0.43
Fuel Efficiency (Combined mpg)	24,133	29.35	3.31
Vehicle Weight (lbs)	24,133	1,360	154
Panel C: Prefecture Characteristics			
Sex Ratio (Male/Female*100)	292	105.28	4.22
Income (GDP per capita)	254	45,324	29,611
Population (10,000)	254	135.40	152.47
Mean House Price in Jan 2010 (per square meters)	97	7,037	4,556
Paved Road Area (per capita)	253	10.87	7.29
Buses (Per 10,000 people)	254	7.72	7.75
Taxis (Per 10,000 people)	254	22.26	18.20

Notes: Pooled sample spanning 2009-2011. Expenditures are deflated to real 2010 RMB prices using the CPI. Sample excludes individuals earning in excess of 250,000 RMB per year. Source: Panel A and B: Authors' calculations using vehicle loan data detailed in Section 4; MPG information obtained from the Ministry of Industry and Information Technology of China (2014). Panel C: Most variables are obtained from China Economic Information Network (CEInet) Statistics Database. Sex ratios are derived from the provincial statistical yearbooks (2011) and mean house prices are obtained from www.elivecity.cn.

Table 2.2: Mean Differences in Differences of Car Price by Gender, Marital Status, and Sex Ratio

	Men		Women		Difference		Difference in
	Unmarried	Married	Unmarried	Married	Unmarried	Married	Difference
Sex Ratio	(1)	(2)	(3)	(4)	(5)=(1)-(3)	(6)=(2)-(4)	(7)=(5)-(6)
Top Quartile	131,084	142,438	124,059	141,627	7,025	811	6,214
	(47570)	(50763)	(41981)	(53044)			
Bottom Quartile	119,417	130,114	115,302	125,003	4,115	5,111	-996
	(38078)	(44786)	(35841)	(43586)			
Diff-in-Diff-in-Diff							7,210

Notes: Sources and sample as described in Table 1. Standard deviations are in parenthesis.

Table 2.3: The Impact of the Sex Ratio on the Automobile Purchase Price of Unmarried Males in China

	Dependent Variable: Automobile Sales Price					Probit: Luxury
	1	2	3	4	5	6
Male	36,586**	22,608	23,510	24,749*	29,546**	0.291***
	(16,194)	(16,433)	(16,586)	(13,530)	(14,436)	(0.088)
Unmarried	34,481**	38,869***	37,044***	36,511*	36,289*	0.406**
	(15,212)	(14,662)	(14,334)	(20,165)	(20,253)	(0.206)
Sex Ratio	1,031***	849***	881***	600***	369**	0.004**
	(137)	(147)	(196)	(126)	(149)	(0.001)
Unmarried *Sexratio	-432***	-497***	-479***	-473**	-469**	-0.004***
	(144)	(134)	(132)	(193)	(194)	(0.001)
Unmarried*Male	-42,378**	-42,168**	-36,111*	-37,093	-40,759	-0.247***
	(17,829)	(19,622)	(19,243)	(26,863)	(27,978)	(0.051)
Sexratio*Male	-322**	-285*	-294*	-304**	-341**	-0.004***
	(147)	(148)	(150)	(129)	(139)	(0.001)
Sexratio*Male *Unmarried	401**	505***	443**	451*	480*	0.005***
	(172)	(183)	(180)	(258)	(269)	(0.002)
Mean Prefecture House price					-0.33*	0.000
					(0.17)	(0.000)
Borrower Characteristics	N	Y	Y	Y	Y	Y
Prefecture Characteristics	N	N	Y	Y	Y	Y
Province Fixed Effects	N	N	N	Y	Y	Y
Mean of Car Price	130,771	130,771	130,806	130,806	128,869	0.239
Number of Obs	24,133	24,133	23,520	23,520	14,929	14,929
R2	0.017	0.143	0.144	0.150	0.156	0.105

Note: Sources and sample as described in Table 1. Borrower characteristics include age, age squared, earnings, and dummies for education level. Prefecture characteristics include income per capita, population, paved road area, buses, and taxis per capita. Standard errors are clustered at the province level in specifications (1)-(3) and reported in parenthesis, while columns (4) and (5) report Huber-White robust standard errors in parenthesis. Column (6) reports marginal effects from a probit regression with a dependent variable of an indicator for luxury car status. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.4: Excluding Cities with Ownership Restrictions

	Dependent Variable: Automobile Sales Price				
	1	2	3	4	5
Male	36,863.57**	19,449.57	19,755.79	21,077.00	26,956.20**
	(15,905.76)	(15,815.61)	(15,904.49)	(15,675.42)	(12,839.15)
Unmarried	35,602.73**	37,400.47**	34,887.62**	33,604.08**	34,271.76**
	(14,375.00)	(15,525.67)	(14,767.52)	(14,115.36)	(14,243.03)
Sex Ratio	1,021.11***	795.33***	822.79***	579.98**	319.05*
	(139.86)	(154.33)	(199.81)	(242.01)	(171.96)
Unmarried *Sexratio	-440.57***	-479.42***	-454.68***	-442.06***	-445.03***
	(136.90)	(143.60)	(137.31)	(131.49)	(131.59)
Unmarried *Male	-41,921.58**	-40,664.60**	-37,062.16*	-36,903.49*	-37,800.37**
	(17,755.54)	(19,185.87)	(20,038.23)	(19,975.29)	(17,537.81)
Sexratio*Male	-325.75**	-256.32*	-259.14*	-270.43*	-316.61***
	(144.68)	(143.84)	(144.78)	(142.46)	(113.28)
Sexratio*Male* Unmarried	398.55**	489.82**	452.68**	449.11**	450.79***
	(171.29)	(179.28)	(188.39)	(188.93)	(157.97)
Borrower Characteristics	N	Y	Y	Y	Y
Prefecture Characteristics	N	N	Y	Y	Y
Province Fixed Effects	N	N	N	Y	Y
Housing Price Controls	N	N	N	N	Y
Mean of Car Price	130,946	130,946	130,987	130,987	129,052
Number of Obs	23,339	23,339	22,726	22,726	14,135
R2	0.017	0.135	0.137	0.143	0.149

Notes: FGLS estimates. Sources, sample, and controls as described in Table 4. Standard errors are clustered at the province level in specifications (1)-(3) and reported in parenthesis, while columns (4) and (5) report Huber-White robust standard errors in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.5: FGLS Estimates (Variance as a Function of Earnings)

	Dependent Variable: Automobile Sales Price				
	1	2	3	4	5
Male	34,107.68**	17,855.65	17,966.41	19,163.70	23,896.85*
	(15,620.13)	(14,978.61)	(14,971.25)	(14,766.65)	(11,976.82)
Unmarried	32,572.44**	34,448.65**	31,942.09**	31,007.41**	30,952.79**
	(15,168.15)	(14,660.48)	(13,927.77)	(13,433.41)	(13,353.83)
Sex Ratio	997.76***	786.94***	811.81***	564.03**	295.41*
	(134.73)	(146.99)	(186.80)	(238.57)	(171.11)
Unmarried *Sexratio	-413.07***	-453.02***	-428.42***	-418.67***	-415.72***
	(144.78)	(135.53)	(129.54)	(125.18)	(123.11)
Unmarried*Male	-42,692.22**	-41,587.42**	-37,304.11*	-36,907.21*	-37,159.15**
	(17,202.17)	(18,340.38)	(18,947.05)	(18,983.67)	(16,071.10)
Sexratio*Male	-301.05**	-241.49*	-242.67*	-252.89*	-289.10**
	(142.83)	(136.35)	(136.29)	(134.19)	(105.18)
Sexratio*Male *Unmarried	405.43**	497.73***	454.43**	448.63**	444.43***
	(166.50)	(171.84)	(178.63)	(180.23)	(145.43)
Mean Prefecture Houseprice					-0.29
					(0.19)
Borrower Characteristics	N	Y	Y	Y	Y
Prefecture Characteristics	N	N	Y	Y	Y
Province Fixed Effects	N	N	N	Y	Y
Mean of Car Price	130,771	130,771	130,806	130,806	128,869
Number of Obs	24,133	24,133	23,520	23,520	14,929
R2	0.017	0.143	0.144	0.150	0.156

Notes: Sources, sample, and controls as described in Table 3. Standard errors are clustered at the province level in specifications (1)-(3) and reported in parenthesis, while columns (4) and (5) report Huber-White robust standard errors in parenthesis.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.6: Robustness Checks Using Age Adjusted Sex Ratios

	Dependent Variable: Automobile Sales Price					
	OLS			FGLS		
	(1) Prefecture Overall sex ratio	(2) Prefecture 20-49 years old	(3) Province 20-49 years old	(4) Prefecture Overall sex ratio	(5) Prefecture 20-49 years old	(6) Province 20-49 years old
Male	23,510	12205	47436**	17,966	5603	46833*
	(16,586)	(9258)	(21707)	(14,971)	(12751)	(27504)
Unmarried	37,044***	21952*	50196*	31,942**	13150	38524
	(14,334)	(12892)	(25494)	(13,927)	(15338)	(25921)
Sex Ratio	881***	554***	636**	811***	431**	623*
	(196)	(150)	(284)	(186)	(159)	(333)
Unmarried *Sexratio	-479***	-371***	-650***	-428***	-276*	-526**
	(132)	(118)	(243)	(129)	(137)	(243)
Unmarried*Male	-36,111*	-20412	-46037	-37,304*	-18003	-43900
	(19,243)	(14790)	(30926)	(18,947)	(16791)	(27191)
Sexratio*Male	-294*	-210**	-557***	-242*	-149	-552**
	(150)	(84)	(208)	(136)	(116)	(266)
Sexratio*Male *Unmarried	443**	322**	576*	454**	294*	550**
	(180)	(137)	(299)	(178)	(155)	(262)
Borrower Characteristics	Y	Y	Y	Y	Y	Y
Prefecture Characteristics	Y	Y	Y	Y	Y	Y
Province Fixed Effects	N	N	N	N	N	N
Housing Price Controls	N	N	N	N	N	N
Mean of Car Price	130,771	130,771	130,771	130,771	130,771	130,771
Number of Obs	23,520	23,520	23,520	23,520	23,520	23,520
R2	0.131	0.132	0.107	0.108	0.089	0.089

Notes: Sources and sample as described in Table 2. Borrower characteristics include age, age squared, earnings, and dummies for education levels. Prefecture characteristics include income per capita, population, paved road area, buses, and taxis per capita. Standard errors are clustered at the province level and reported in parenthesis. Column (1) and (4) are copied from column (3) of Table 3 and Table 4 for comparison. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.7: Exploring Earnings Heterogeneity

	Dependent Variable: Automobile Sales Price				
	1	2	3	4	5
1st Quartile (0, 60,000)	1,033***	920***	794***	769***	717**
	(251)	(255)	(241)	(260)	(277)
2nd Quartile (60,000, 90,000)	-283	-239	-149	-94	-71
	(322)	(326)	(268)	(266)	(204)
3rd Quartile (90,000, 12,000)	710	754*	658	739*	487
	(466)	(404)	(425)	(426)	(288)
4th Quartile (12,000, 25,000)	499	550	511	479	954*
	(609)	(513)	(512)	(515)	(538)
Borrower Characteristics	N	Y	Y	Y	Y
Prefecture Characteristics	N	N	Y	Y	Y
Province Fixed Effects	N	N	N	Y	Y
Housing Price Controls	N	N	N	N	Y

Notes: Sources, sample, and controls as described in Table 4. FGLS estimates of the triple interaction coefficient from Table 4. Overall sample is 24,133 split into quartiles. Standard errors are clustered at the province level in specifications (1)-(3) and reported in parenthesis, while columns (4) and (5) report Huber-White robust standard errors in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.8: Exploring Regional Sex Ratio Heterogeneity

	Dependent Variable: Automobile Sales Price				
	1	2	3	4	5
1st Quartile (89.67, 102.74)	-368	305	330	377	651
	(542)	(440)	(418)	(364)	(490)
2nd Quartile (102.74, 105.14)	2,098	5,442	6,481*	6,654*	5,015
	(3,640)	(3,526)	(3,487)	(3,465)	(2,990)
3rd Quartile (105.14, 107.75)	282	-2,536	-1,491	-1,016	-4,405
	(4,661)	(2,556)	(2,504)	(2,852)	(2,950)
4th Quartile (107.75, 132.40)	724***	403**	374**	353**	280
	(189)	(150)	(161)	(147)	(210)
Borrower Characteristics	N	Y	Y	Y	Y
Prefecture Characteristics	N	N	Y	Y	Y
Province Fixed Effects	N	N	N	Y	Y
Housing Price Controls	N	N	N	N	Y

Notes: Sources, sample, and controls as described in Table 4. FGLS estimates of the triple interaction coefficient from Table 4. Overall sample is 24,133 split into quartiles. Standard errors are clustered at the province level in specifications (1)-(3) and reported in parenthesis, while columns (4) and (5) report Huber-White robust standard errors in parenthesis.

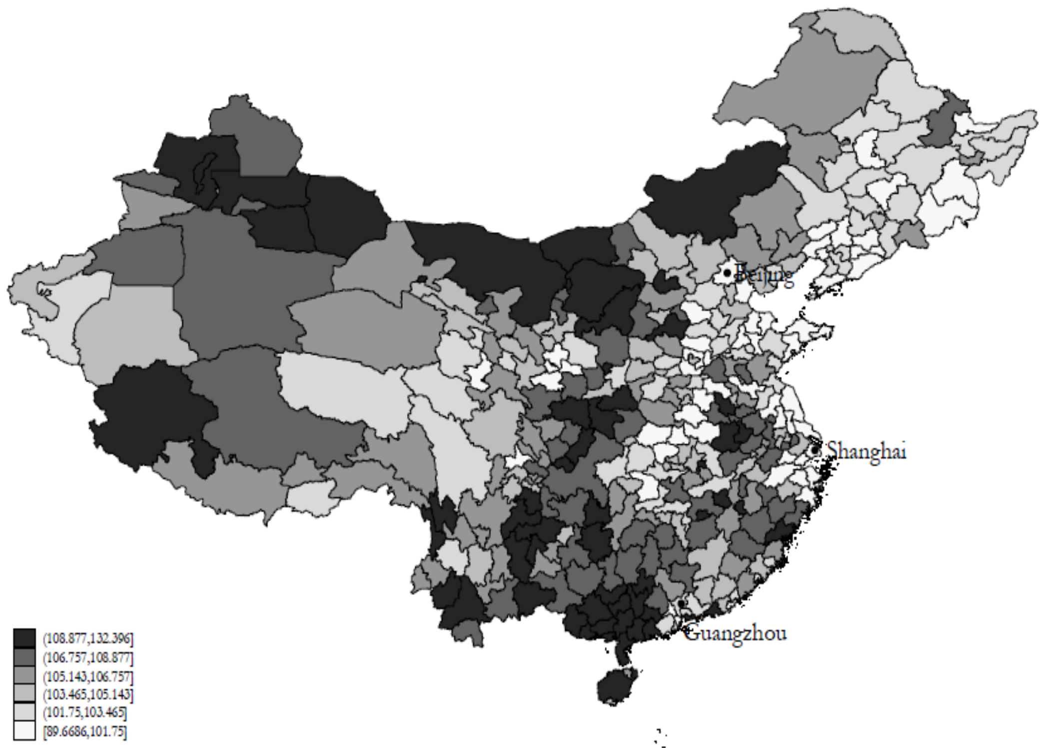
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.9: Externalities Associated with Luxury Purchases

Dependent Variable	1	2	3	4	5
Fuel Efficiency	-0.047***	-0.038***	-0.030***	-0.030***	-0.032**
	(0.014)	(0.012)	(0.010)	(0.010)	(0.013)
Vehicle Weight	2.44***	1.67**	1.54**	1.55**	1.76***
	(0.67)	(0.60)	(0.60)	(0.61)	(0.54)
Borrower Characteristics	N	Y	Y	Y	Y
Prefecture Characteristics	N	N	Y	Y	Y
Province Fixed Effects	N	N	N	Y	Y
Housing Price Controls	N	N	N	N	Y

Notes: Sources, sample, and controls as described in Table 4. FGLS estimates of the triple interaction coefficient from Table 6 for the fourth quartile of sex ratio. Samples range from 5,973 to 5,703 observations except for column (6) which has 2,337 observations. Standard errors are clustered at the province level in specifications (1)-(3) and reported in parenthesis, while columns (4) and (5) report Huber-White robust standard errors in parenthesis.
*** p<0.01, ** p<0.05, * p<0.1

Figure 2.1: Sex Ratios by Prefecture



Chapter 3: Equity Market Liberalization and Economic Growth Revisited: New Data and New Methods

3.1. Introduction

Equity market liberalization, also known as stock market liberalization, enables foreign investors to trade in a country's domestic equity securities market, and is a form of financial liberalization.[20] Growth theory does not provide a clear prediction concerning how liberalization should impact economic growth. On the one hand, there are many channels through which equity market liberalization may contribute to growth. The most direct channel is that liberalization attracts foreign funding for domestic companies. In addition, liberalization could also facilitate financial development (Fuchs-Schundeln *et al*., 2003), reduce capital cost (Henry, 2000a; Bekaert and Harvey, 2000), improve risk-sharing (Iwata and Wu, 2009), boost private investment (Henry, 2000b), enhance export (Manova, 2008), improve corporate control (Levine, 2002), and ameliorate firm-level operating performance (Mitton, 2006). There are other benefits associated with liberalization such as higher equity prices (Henry, 2000a), higher turnover ratio and better compatibility of incentives of shareholders and managers (Fuchs-Schundeln *et al*., 2003). On the other hand, liberalization could generate extra risk, induce macroeconomic volatility, or increase the frequency of crises. Moreover, in the presence of information asymmetries, foreign capital may be invested inefficiently,

[1]Broadly speaking, financial liberalization consists of equity market liberalization, capital account liberalization and banking sector liberalization.

leading to a slower output growth (Stiglitz, 2000). Therefore, the net impact of liberalization on growth is ambiguous in theory and merits empirical evaluation.

Previous empirical research has failed to arrive at a consensus, finding somewhat mixed results.[21] For instance, Bekaert et al. (2005) document a 1 percent increase in annual real growth rate associated with liberalization and show that the first five years after liberalization account for most of the impact. Fuchs-Schundeln et al. (2003) find a 4 percent cumulative short-term growth effect within 4 years after liberalization as well as a 0.4 percent per year permanent growth effect. Bonfiglioli et al. (2004) compare the impact of equity market liberalization with that of capital account openness and show that the former affects growth directly while the latter impacts growth by altering the financial depth. In contrast, Ben Naceur et al. (2008) inspect cases in the Middle East and North Africa (MENA) but fail to find any significant relationship between liberalization and growth. Similarly, Ben Gamra (2009) studies East Asian countries and reports a weak negative effect.

In this paper, we start out by replicating the key results in Bekaert et al. (2005), a highly influential and comprehensive study in this topic. Then we contribute to the literature by carrying out two extensions. Probably the most important extension is the expansion of the database. While previous studies provide great insight into the growth effect of liberalization, to the best of our knowledge, all of them utilize data on stock market liberalization prior to 2000. We close this gap by enlarging the data through 2011, which leads to the inclusion of 25 additional liberalization cases that are not

[21] A large body of existing research has attempted to empirically evaluate the relationship between overall financial market liberalization and economic growth, but finds mixed evidence. Detailed surveys can be found in Hermes and Lensink (2005) and Kose *et al*. (2009), and a meta-analysis in Bumann *et al*. (2013). Most of these studies are narrowly focused on the impact of capital account liberalization, leaving only a small portion of them examining the growth effect of equity market liberalization.

considered in previous research. Among these liberalizations, some emerged after 1999, others are "old" cases that occurred before 2000 but have not been identified or have been neglected by the existing literature. We decide to include all countries that opened up their stock markets between 1980 and 2011 for the analysis. We also perform robustness checks by excluding countries with special circumstances surrounding liberalization.

Our second extension is the use of propensity score matching (PSM) to address endogenous policy selection. Most papers that study the effects of liberalization have employed panel data estimations via OLS, seemingly unrelated regression (SUR), fixed effects, or GMM. A common limitation of these studies is that they rely on the assumption that liberalizations are exogenous, ignoring the possibility of endogenous policy selection.[22] This drawback could bias the estimates of the treatment effect if, on average, countries that adopted liberalization policies are fundamentally different from countries that did not, in terms of the determinants of economic growth. Thus, it is of great importance to construct a valid counterfactual in order to derive an unbiased causal impact of liberalization.

To accomplish this, we employ propensity score matching methods introduced by Rosenbaum and Rubin (1983, 1985). We further combine matching with a hazard model to take account of the temporal variation in our data. As a first step, a Cox proportional hazard model is applied to estimate the probabilities (propensity scores) of liberalization for each country-year. The scores are then used to find a valid comparison group in the matching process. Finally, we use difference-in-difference to find the one-

[22]Bakeart *et al*. (2005) attempt to address the endogeneity by controlling for exogenous growth opportunities a country faces.

year average through ten-year average growth effects associated with equity market liberalization. We find that correcting for the selection bias leads to more persistent growth effects than previous research has documented.

The remainder of the paper proceeds as follows. Section 2 discusses the data. Section 3 replicates Bekaert *et al*. (2005) and applies their methods on our expanded dataset. Section 4 conducts the survival analysis and propensity score matching. Section 5 summarizes the paper.

3.2. Data

Our sample covers 166 countries/regions that had not opened their stock markets as of 1980. Among them, 65 economies liberalized during the period 1980-2011. This treatment group is composed of 10 countries from Latin America and the Caribbean, 10 from Sub-Saharan Africa, 11 from the Middle East and North Africa, 14 from Asia and Oceania, and 20 from Europe. The most important variable in this study is the date of equity market liberalization. Three competing methods have been employed in the literature to identify the liberalization dates. Official decree date, which is the date of the official announcement of liberalization, is a natural choice and is probably the most frequently used method. The second type of the liberalization date is the earliest date of the following three events: official decree announcement, first American Depositary Receipt (ADR) announcement, and first country fund launch.[23] Finally, if researchers are interested in differentiating levels of liberalization, they can use liberalization

[23]ADR is a security representing shares in a foreign stock that is traded in U.S. financial markets and is denominated in U.S. dollars. A country fund is usually launched by investment companies, and makes investments in equity shares from a particular foreign country.

intensity indices, such as the share of domestic equity securities that foreign investors can purchase.

We adopt the official decree date as our liberalization date. The data for the period 1980-1997 mainly come from Bekaert *et al*. (2005) and Fuchs-Schundeln *et al*. (2003).[24] We enlarge the database by including cases between 1998 and 2011, primarily based on Standard & Poor's Global Stock Market Factbook, annual Investment Climate Statements from U.S. Department of State, *A Chronology of Important Financial, Economic and Political Events in Emerging Markets* compiled by Bekaert and Harvey, and a few other papers and books. We also incorporate liberalizing countries between 1980 and 1997 that have failed to be recognized by existing literature.[25]

The liberalization dates used in this study are presented in Table 3.1. We further provide the reasons for our date choices for the cases that previous papers did not include or the cases where our liberalization dates differ from those used in the past research. Interestingly, all liberalizations occurred prior to 2005.

Finally, some special cases warrant additional attention. Estonia, Latvia, Lithuania, Russia, and Ukraine emerged from the collapse of the Soviet Union in 1991; Croatia and Slovenia became independent after the breakup of Yugoslavia in 1991 and Czech and Slovakia emerged from the dissolution of Czechoslovakia in 1993. All of them liberalized soon after independence, which could have been followed by social and economic turmoil. The second special case is China. Although China established a B-

[24] Whenever these two studies assign different dates for a country, we examine their reasons and determine the dates for our paper.

[25] Actually, we find twelve more liberalizing countries/regions during the period 1980-1997 that previous papers did not consider in their analyses: Chinese Taipei 1991, Russia 1991, Namibia 1992, Poland 1994, Lebanon 1994, Zambia 1994, Czech 1995, Hungary 1995, Estonia 1996, Latvia 1996, Ukraine 1996 and Cyprus 1997.

share market open to foreign investors in 1992, we do not treat that year as the liberalization date because the B-share market is quite small and illiquid compared with the A-share market, which is the principal stock market in China. Our preferred date is 2003 when the A-share market opened to Qualified Foreign Institutional Investors (QFII). In the following analysis, robustness checks will be conducted without these countries.

The remaining data is associated with country-level characteristics and primarily comes from the World Bank Development Indicators (WDI). Information on Taiwan is not available from the WDI. We obtain its data from other sources such as Penn World Table 7.1, Ministry of Education of Taiwan, and Ministry of the Interior of Taiwan. Table 3.2 lists all the variables used in this study and their sources.

3.3. Replication

We first replicate Bekaert *et al*. (2005)'s analysis concerning the effects of equity market liberalization on economic growth.[26] Table 3.3 presents a summary of the results. Following their paper, we mainly adopt two different samples. One of them only includes the countries that liberalized during 1980-1997. We calculate the difference between the post-liberalization and pre-liberalization 5-year average growth rates of real per capita GDP in column (1). We then run three regressions with fixed effects, time effects and both in column (2) through (4) respectively. In each regression, the

[26]Bekaert *et al*. (2005) also perform many robustness checks, compare equity market liberalization and capital account liberalization, explore the heterogeneity of the growth effect, attempt to address the endogeneity issue and in some regressions add the covariates one by one and eventually all together. We do not aim to replicate all their results. Instead, we only focus on their estimation of the growth effect, with a full set of independent variables.

dependent variable is the annual growth rate of real per capita GDP and the independent variable is a dummy that takes a value of one if the stock market is liberalized and zero otherwise. There are no other control variables used.

The second sample depends on data availability. Growth rates of GDP per capita, the liberalization indicator, macroeconomic and demographic variables must all be available for the countries in this sample.[27] It consists of both liberalizing countries and countries that never liberalized during 1980-1997. Two regressions are conducted using this sample. Column (5) reports the growth effect estimated in the OLS regressions. The dependent variable is non-overlapping five-year average growth rate. Three separate OLS regressions are carried out over three different periods (1981-1995, 1982-1996 and 1983-1997), and the simple average of the three coefficients are presented. Finally, column (6) shows the results of a generalized method of moments described in detail in Bekaert *et al*. (2001, pp. 472-479). The dependent variable is overlapping five-year average growth rate. This GMM technique has several advantages. It allows researchers to better utilize the temporal components of the data and can adjust the standard errors for using overlapping observations. Additionally, with different weighting matrices, it can address heteroscedasticity across countries, heteroscedasticity across time and/or seemingly unrelated regression effects.

As is shown in the first two rows of Table 3.3, while we are unable to precisely replicate Bekaert *et al*. (2005)'s findings, our results are close, with the same signs and

[27] More specifically, the independent variables include the natural logarithm of GDP in 1980, government spending as a share of GDP, secondary school enrollment, population growth rate and the natural logarithm of the life expectancy.

similar coefficients.[28] We then extend the analysis by using the updated data in the third through fifth rows and find that the growth effect of liberalization remains strong and significant, although OLS and GMM imply a slightly weaker effect. The results are also quite robust to excluding China and countries that emerged from the dissolutions of the Soviet Union, Yugoslavia, and Czechoslovakia.

3.4. Addressing the Self-Selection Issue Using Matching

Most previous research concerning the effects of equity market liberalization on economic outcomes treats liberalization decisions as if they are exogenous. However, it could be the case that countries select themselves into the liberalizing group based on a variety of economic factors. If this is true, the liberalizing countries may be fundamentally different than the rest of the countries such that the estimated average treatment effects generated by conventional regressions are biased. This section, therefore, is intended to address endogenous policy selection using propensity score matching.[29]

3.4.1. Survival Analysis

The first step of propensity score matching is to generate the propensity score using a probability model. A majority of studies employs either a probit or logit regression. However, since countries carry out liberalizations in different years, these two models

[28]We collected data from the same source (mostly the WDI) as Bekaert *et al*. (2005) did, but still cannot precisely replicate BHL's results. An important reason is that data from the WDI may have changed over time. For instance, we cannot find any observations on Jamaica's real GDP per capita from the current WDI dataset and have to exclude it from our analysis while Bekaert *et al*. (2005) use it as a liberalization case.

[29]In order for our findings to be comparable with the existing literature, we use the same sample as in Bekaert *et al*. (2005) in our matching process. The results for the full sample are available in Appendix A.

may be unsuited for the policy under study as they fail to take account of the temporal feature of the treatment. Given this time-dependency and the fact that liberalization is rarely repealed, a hazard model is a good fit. Specifically, we use a Cox proportional hazard regression to model the liberalization decisions. The Cox model is convenient, reasonable, and most importantly valid to be used in PSM (see Lu (2005) for a proof).

It is critical to correctly assign the right-hand-side variables for the hazard model, not only because it is essential to understand what factors are driving the liberalization reforms, but the policy predictors are exactly the covariates intended to be matched upon. Many studies employing PSM include as predictors only the variables that simultaneously affect policy adoption and the outcome variable. A few others exclusively control for the determinants of the former. These exercises could be inappropriate, as Cuong (2013) demonstrates with Monte Carlo simulations that when estimating the average treatment effect on the treated (ATT), more efficiency can be gained if all the determinants of the outcome variable are used in the matching process, even if they do not influence the policy adoption. Rubin and Tomas (1996) also point out that a variable should not be excluded from the analysis unless there is a consensus that the variable is either irrelevant to the outcome or it is not a proper covariate. Intuitively, if we omit important variables that impact the outcome, the matched samples are likely to be unbalanced with respect to these ignored variables, which in turn can contribute to the difference in the outcome variable between the treatment and the control groups. After all, the ultimate goal of PSM is to allow fewer confounders (factors that affect the outcome) to interfere with the results. Therefore, we will include both determinants of policy adoption and determinants of GDP growth as covariates in the estimation of the propensity scores. Finally, we lag all the covariates by one year to avoid contemporaneous endogeneity and to take into account the fact that a decision to liberalize is made before the policy is implemented.

We mainly consider two sets of covariates. One of them is the policy predictors. Kaya *et al*. (2012) provide valuable guidance regarding the economic factors that drive the equity market liberalization decisions. They find that industrialization level, financial development, legal origin (a proxy for the quality of investor protection), government expenditure (a proxy for the level of government involvement in the economy), and received foreign aid are significantly correlated with the liberalization decisions. So these variables will be part of our right-hand-side variables. In particular, we use credit to the private sector rather than stock market turnover as a measure of financial development because the data of the latter are sparse. With regard to the legal origins, we use a dummy variable that indicates a legal system based on common law, which is consistent with the literature.

The second set of covariates is the determinants of GDP growth. There is a rich pool of them and our choice of the covariates is primarily guided by Grier and Grier (2007). Investment, education, and technology are probably the most classical ones as they play important roles in nearly every neoclassical growth model. Investment was once deemed as a panacea for growth; education determines the level of human capital and the overall labor effectiveness; and technological advance is regarded as a major source of growth over the long run. We use investment as a share of GDP and primary school enrollment rate to represent investment and education, respectively. There are two variables, namely the expenditure on R&D and the number of patent applications that the literature has frequently used to reflect technological progress. We choose the latter to take advantage of more observations.

In addition to the classical ones, institution, inflation and initial GDP per capita are often deemed as determinants of growth as well. For example, in a highly influential paper, Acemoglu *et al*. (2001) show that institutions have large effects on the long-term economic performance. We use political constraints on the executive as a measure of institutional quality. Inflation can also affect growth. Evidence has been found that both inflation per se and inflation uncertainty are detrimental to growth (Barro, 1996; Grier and Grier, 2006). We use GDP deflator to represent inflation. According to the convergence hypothesis in economics, countries with a lower starting level of real per capita GDP tend to grow at a higher rate than richer economies. While the validity of this notion has long been a dispute, empirical studies routinely include the initial GDP level as a growth predictor. Finally, to take account of current account and capital account liberalization that may go in tandem with stock market liberalization, we control for openness to trade as well as the Chinn-Ito index, a de jure measure of financial openness, in our regression.

Ideally, every covariate described above can be used in the survival analysis (and subsequently controlled for in matching) such that the results will be least biased. However, due to the data constraints, more covariates come at the cost of smaller samples, which reduce the precision of the estimators. In Table 3.4, we report the results of the Cox hazard regressions. Column (1) has the most parsimonious specification but take into account 40 liberalizing countries. Column (3) has a full set of covariates but suffers from a small sample bias. In the face of the tradeoff, we also use a third specification in column (2), which considers most covariates (11 out of 13) and still retains 36 treated countries.[30] Despite the different explanatory variables included in the regressions, our results largely confirm Kaya *et al*. (2012)'s findings. Countries with common law legal origins are more likely to liberalize as they tend to provide better protection of investors' rights. Government size is inversely related to the probability of liberalization because a big government that uses political power to allocate resources is less likely to approve liberalization. Better financial development drives liberalization given that a sound financial system would be more capable of absorbing and utilizing foreign capital. The level of industrialization is also positively associated with liberalization since an agricultural country is less likely to open its stock market. Finally, more received financial aid tends to stimulate liberalization as it helps strengthen the reforms in policies and institution (Kaya *et al*., 2012).

One might be concerned that only a few covariates exert a significant impact on the dependent variable. However, since the primary purpose of matching is to balance the distribution of these covariates such that the treatment group and the control group are similar and comparable, rather than to accurately estimate each covariate's effect on policy adoption, all the covariates in the model will be retained.

3.4.2. Matching

In this subsection, we match the treated country-years to a control group based on the fitted values generated in the hazard regressions. It is convenient to use the linear portion of hazard (i.e. the estimated value of $X\beta$ in the Cox hazards model) as the

[30]Actually, the results are fairly robust to a diversity of specifications, not limited to the three that we report.

propensity score (Lu, 2005; Ufier, 2014). Figure 3.1 shows the distribution of propensity scores for both the treated and control groups. To make sure that the treated and control units share a common support, we drop treated observations whose propensity scores are above the maximum or below the minimum of the controls'. Various matching techniques can then be chosen, among which single nearest neighbor matching with replacement is the most commonly used one. In this method, each liberalizing country-year is matched with another country-year that has not liberalized and has the closest propensity score. Other matching methods such as radius matching and kernel matching are also popular. See Rosenbaum and Rubin (1983, 1985) for more details about matching estimator properties and Dehejia and Wahba (2002) for a discussion of choosing appropriate matching methods under different circumstances. In this paper, we employ kernel (normal) matching as our primary method because it uses every potential control country within the common support and that it appears to be more capable of creating control groups with balanced covariates in this study. We will also report other matching results as robustness checks.

Given that the propensity scores are estimated for every country in every year, our matching algorisms allow the treated countries to be matched to controls in different years than the liberalization dates, as long as they closely resemble each other and have similar propensity to liberalize. This tremendously expands the number of potential controls and helps find much better matches on the observed covariates. Each treated country is matched to one or more countries in the control pool, which contains countries that are exposed to treatment in later years as well as those that are never treated. As we seek to assess the average growth effects over time for 10 years, the controls must not be treated within 10 years. Therefore, countries 10 years or less before they liberalize should also be excluded from the control pool.

After matching, it is important to verify that the differences in observable confounders between the treatment group and the control group have been eliminated or at least reduced to ensure an unbiased estimate. Table 3.5 shows the result of kernel (normal) matching by comparing the means of covariates used in specification (3) of Table 3.3.[31] It turns out that the covariate balance is well achieved: Prior to matching, 8 covariates show a statistically significant difference at the 10% level, and all of them disappear after matching.

3.4.3. Results

Having obtained properly matched treated group and control group, one may be tempted to directly compare the average post-treatment GDP growth rate between these two groups. However, it might be the case that the treated countries have been experiencing a higher growth rate prior to liberalization and would have grown at a higher rate had liberalization not taken place. This can result in a false estimate of a positive effect of liberalization on growth even if it has no effect at all. To address the issue, we adopt a difference-in-difference analysis by comparing the change in the growth rate (i.e. the dependent variable is N-year average post-treatment growth rate less N-year average pre-treatment growth rate, N being 1, 2,…,10) between the treated

[31]We also check covariate balances for other hazard model specifications and for other matching techniques. When fewer covariates are included (i.e. using column 1 or 2 in table 3), balance is almost always well achieved. When column 3 of table 3 is used, 13 covariates need to be balanced at a time. Occasionally one or two unbalanced covariates may show up. In this situation, we modify the hazard regressions by adding higher order terms of some covariates and redo matching. In essence, this exercise alters the weights of the covariates in matching. Refer to Dehejia and Wahba (2002) for details.

and control groups.[32, 33, 34] Another benefit associated with the difference-in-difference matching is that it differences out the time-invariant unobservable confounders which matching could not control for. The results are reported in Table 3.6: Equity market liberalization stimulates growth substantially, which is robust to various matching methods. When the smallest sample (panel 1) is used, liberalizing countries enjoy a huge bonus growth rate immediately (more than 3%), and the effect remains about 1% throughout. When the sample size is larger (panel 2 and 3), we obtain a higher statistical significance. There is strong evidence of a substantial gain (about 1.5%-3%) in the short-term and weak evidence of an appreciable gain over the long run: most estimates for 9-year or 10-year average growth effects are above 0.9% with the statistical significance below or close to the 10% level. This estimate, 0.9% annually for a whole decade, should be big enough to arouse the interests of political leaders whose countries have yet to liberalize their stock markets. Compared with the literature, our findings are largely in accordance with the results in Bekaert *et al*. (2005) and Fuchs-Schundeln *et al*. (2003), who find an average growth effect of about 1% after liberalization. But our estimates indicate a stronger effect in the first two years after liberalization and hint at a more sustainable effect that persists for 10 years. We also conduct matching analysis for our expanded data. The results are presented in Table 3.7. The addition of the new cases

[32] The post-treatment growth rates do not include the rate in the liberalization year because liberalization could occur at the end of the year. It is inappropriate to view that year as post-treatment and the estimated effect in that year is expected to be misleadingly small.

[33] Another way to deal with the reverse causation issue is to include pre-treatment growth rate as a variable to be matched upon. The problem with this method is, in most cases, even though the pre-treatment growth rates are balanced between the treated group and the control group (which means they are not statistically significantly different), their difference can still be large enough to affect the results. For example, 3.89% and 4.85% are statistically different at the 30.1% significance level, which passes the covariance balance test. But their difference almost reaches 1%.

[34] Note that a positive effect could either because treated countries grow faster after liberalization or because control countries grow slower or both. This explains some of the surprisingly large estimates we obtain.

seems to diminish the immediate effect. But the results are generally still supportive of the conclusion that liberalization promotes economic growth by about one percentage point annually.

3.4.4. Placebo Test

It is of interest to know whether the results we obtain above truly come from liberalization rather than due to randomness. A placebo test is carried out where we apply the same matching techniques to the liberalizing countries, but the liberalization dates are artificially moved to 10 years prior to the true dates. We then examine the one-year through ten-year average growth effects associated with these imaginary policy interventions. Since the liberalizations did not actually occur, we expect the estimates of the growth effects to be statistically indistinguishable from zero. Table 3.8 confirms our conjecture that the results can be either positive or negative and most importantly, none of them are significant at the 10% level.

3.4.5. Does matching make a difference?

This subsection demonstrates why matching is important by comparing our matching results with the results obtained from panel regressions. Following Fuchs-Schundeln *et al*. (2003), we specify a standard growth regression that is commonly used in the literature:

(1) $$Growth_{it}=\alpha_i + year_t + \beta_0 lib0_{it} + \beta_1 lib1_{it} + \beta_2 lib2_{it} + \beta_3 lib3_{it} + \beta_4 lib4_{it} + \beta_5 lib5_{it}$$

$$+ \beta_6 lib6_{it} + \beta_7 lib7_{it} + \beta_8 lib8_{it} + \beta_9 lib9_{it} + \beta_{10} lib10_{it} + \gamma' X_{it} + \varepsilon_{it}.$$

where the dependent variable is the real GDP growth rate per capita of country i in year t. Dummy lib0 takes the value of one in the year of liberalization and zero otherwise. Similarly, lib1-lib10 refer to the dummies indicating the years after liberalization. X is a vector of independent variables that we have used in the matching process. In addition, we control for country fixed effect "α_i" to alleviate potential omitted variable bias and year fixed effect "$year_t$" to account for possible worldwide growth trend or common external shocks that impact every country. For comparison, we also include pooled OLS and country fixed effect regression.

The results are reported in Table 3.9. Consistent with Bekaert *et al*. (2005) and Fuchs-Schundeln *et al*. (2003), these regressions capture a 1% to 2% annual growth effect for three or four post-treatment years. Beyond this time window, the coefficients are mostly small, insignificant, and even negative. In order to make them comparable with the matching results, we calculate the N-year average effect, N=1,…,10, using the following simple algorithm: N-year average effect=$\frac{\sum_1^N \mathrm{Libi}}{N}$, where Lib_i is the fitted value in each regression. Table 3.10 shows the results. The second column represents the estimates from the kernel (normal) matching, which is copied from the second column of Table 3.6, and the remaining columns represent results from panel regressions. We find two important patterns: first, all regressions tend to underestimate the immediate growth effect of liberalization. While matching reveals a 1.5% to 4% growth spurt in the year after liberalization, estimates from regressions are all below 1.5%. Second, regressions fail to recognize the persistence of the growth effect. Their results are typically much smaller than the matching results. In sum, matching does make a difference.

5. Conclusion

The past three decades have witnessed a large number of equity market liberalizations. The existing literature finds mixed evidence on the growth effect associated with liberalization. However, none of them has studied liberalization cases after 1999. In this study, we augment the analysis by extending the liberalization cases through 2011 and including the cases they have overlooked. We closely mimic Bekaert *et al*, (2005) and apply their methods on our updated data. Consistent with their results, we find strong evidence that liberalization leads to a sizable increase in annual real economic growth. We also argue that endogenous policy selection is an important issue that conventional regressions are unable to address. Therefore, we use propensity score matching to control for the selection bias and find that the growth effect of liberalization is likely to persist for a whole decade, which is much longer than the literature has documented.

Table 3.1: Countries that Officially Adopted Equity Market Liberalization and Their Starting Years

Country	Bekaert et al. (2005) 1980-1997	Fuchs-Schundeln (2003) 1980-1995	This Paper 1980-2011	Reason for the Dates
Argentina	1989	1989	1989	
Bangladesh	1991		1991	
Botswana	1990		1990	
Brazil	1991	1991	1991	
Chile	1992	1992	1992	
Colombia	1991	1991	1991	
Cote d’Ivoire	1995		1995	
Ecuador	1994		1994	
Egypt	1992	1993	1992	
Ghana	1993		1993	
Greece	1987	1986	1987	
Iceland	1991		1991	
India	1992	1992	1992	
Indonesia	1989	1989	1989	
Jamaica	1991		1991	
Japan	1983	1980	1980	
Jordan	1995	1995	1995	
Kenya	1995		1995	
Korea, Republic of	1992	1992	1992	
Malaysia	1988	1988	1988	
Malta	1992		1992	
Mauritius	1994		1994	
Mexico	1989	1989	1989	
Morocco	1988	1994	1994	
New Zealand	1987		1987	
Nigeria	1995	1995	1995	
Pakistan	1991	1991	1991	
Peru	1992	1991	1991	
Philippines	1991	1991	1991	
Portugal	1986	1986	1986	
Saudi Arabia	1999		1999	
South Africa	1996	1995	1996	
Spain	1985	1985	1985	
Sri Lanka	1991	1990	1990	
Thailand	1987	1987	1987	
Trinidad and Tobago	1997		1997	
Tunisia	1995		1995	
Turkey	1989	1989	1989	
Venezuela	1990	1990	1990	
Zimbabwe	1993	1993	1993	
Bahrain			1999	…loosening foreign investment restrictions on listed equities. Foreigners are entitled to invest in up to 49% of a domestic public-shareholding company's equities, while seven banks are 100% open to foreign investors
Bulgaria			1998	Kouretas (2012)
Cyprus			1997	In February 1997, the government revised its policy on FDI, permitting 100% foreign ownership in certain cases. Regulations on foreign

			portfolio investment in the Cyprus Stock Exchange also have been liberalized.
Hungary		1995	Under the new foreign exchange law, some capital account restrictions were eliminated; foreigners were allowed to buy most Hungarian securities with maturities longer than one year without obtaining permission from the National Bank, and outward equity investment was permitted, provided that an equity share of over 10% is acquired.
Kuwait		2000	The National Assembly ratified the "Indirect Foreign Investment Law" in August 2000, allowing foreigners to own 100 percent of all listed shareholding companies, except banks.
Lebanon		1994	The Investment Development Authority of Lebanon was created to support foreign investors. With very few exceptions, there is no discrimination between national and foreign investments
Oman	1999	1998	Information from Kim et al. (2007), Mansur et al. (2008), Azzam (2013) and Hassan et al. (2003)
Qatar		2003	In December 2003, the Cabinet of Ministers approved a law allowing foreigners to own up to 25 percent of a company listed in the DSM (Doha Securities Market).
Namibia		1992	The Stock Exchanges Control Act was amended. There are no restrictions on foreign investment (although foreigners need special permission if they wish to take over control of a bank).
Poland		1994	Portfolio investment on the Warsaw Stock Exchange was liberalized in September 1994.
Romania		1998	Kouretas (2012)
Taiwan, China		1991	…implementation date of phase two of liberalization plan. Eligible foreign institutional investors may now invest directly in Taiwan securities if they have applied for and received SEC approval as a qualified foreign institutional investor (QFII). Each foreign institution is limited to holding a maximum of 5% of any listed stock and total foreign holdings in any listed companies may not exceed 10%.
United Arab Emirates		2004	Ministry of Economy and Planning rules allow foreign investment up to 49 percent in companies on the stock market; however, company

		by-laws in many cases prohibit or limit foreign ownership.
Vietnam	2003	…as part of its efforts to encourage foreign investment and to promote the development of the infant stock market, the Government issued Decision 146 in July 2003 abolishing the equity limit of a single foreign investor (institutional or individual) in a listed Vietnamese company.
Zambia	1994	A stock exchange was introduced in 1994. Participation on both the bond and stock markets was open to not only residents but non-residents as well.
Other Cases		
China	2003	The A-share market is open to QFII (Qualified Foreign Institutional Investors) in 2003.
Croatia	1998	FDI, inward portfolio investments, and profit transfers abroad are not restricted.
Czech	1995	Most capital transactions were de jure liberalized with the enactment of the new Foreign Exchange Act in September 1995, including short-term portfolio and credit inflows.
Estonia	1996	The Tallin Stock Exchange opened, and foreign investment service firms can operate there.
Latvia	1996	Amendments to the Investment Law passed in 1996 removed virtually all restrictions on foreign investment. Securities markets are regulated by the 1996 Law on Securities and some other laws.
Lithuania	1999	Kouretas (2012)
Russia	1991	The 1991 investment code guaranteed foreign investors rights equal to those enjoyed by Russian investors.
Slovakia	1998	Kouretas (2012)
Slovenia	2001	Bank of Slovenia lifted all restrictions on foreign portfolio investments in the Slovenian capital market.
Ukraine	1996	The Foreign Investment Law passed. It guarantees foreign investors equal treatment with local companies, and the "unhindered transfer" of profits, revenues, and other proceeds in foreign currency after covering taxes and other mandatory payments.

Notes:

[a]The data are mainly from Bekaert *et al*. (2005), Fuchs-Schundeln *et al*. (2003), Global Stock Market Factbook by S&P, Investment Climate Statements from U.S. Department of State and *A Chronology of Important Financial, Economic and Political Events in Emerging Markets* at http://www.duke.edu/~charvey/Country risk/Chronology/.

[b]The data cover period 1980 to 2011, although all cases are before 2005.

[c]For liberalization dates before 1997, if Bekaert *et al*. (2005) and Fuchs-Schundeln *et al*. (2003) use different dates for a country, we examine their reasons and determine the dates for our paper.

[d]Nations and years in the italic font are cases that previous studies on growth effect of liberalization did not include or cases where our liberalization dates differ from those in the existing literature. For these cases, we provide the reasons for our date choices or the data sources.

[e]"Other cases" include (1) Countries that emerged from the dissolution of the Soviet Union in 1991; (2) Croatia and Slovenia, which became independent after the breakup of Yugoslavia in 1991 (3) Czech and Slovakia, which came from the dissolution of Czechoslovakia in 1993; (4) China where foreign access to "A" stocks is still very limited due to the quotas assigning to foreign investors.

Table 3.2: Descriptions of Variables and Data Sources

Official liberalization date	The year when the equity market is officially liberalized. Source: Bekaert et al. (2005), http://www.duke.edu/~charvey/Country risk/Chronology/ and Maria-Lenuta et al. (2011)
Legal Origin	Equals one if the origin of the commercial law of a country is English Common Law, and zero otherwise. Source: La Porta et al. (1999).
Industrialization level	Value added in industry as a share of GDP. Source: World Bank Development Indicators
Government expenditure	General government final consumption as a share of GDP. Source: World Bank Development Indicators
Private credit	Credit to private sector divided by GDP. Source: World Bank Development Indicators
Inflation	Inflation measured by GDP deflator. Source: World Bank Development Indicators
Real Per capita GDP	Per capita GDP (2005 $) Source: World Bank Development Indicators
Real GDP growth rate	Growth rate of real GDP per capita Source: World Bank Development Indicators
Financial aid	Development assistance and financial aid received ($) Source: World Bank Development Indicators
Primary school enrollment	% of gross primary school enrollment Source: World Bank Development Indicators
Population growth rate	Growth rate of total population Source: World Bank Development Indicators
Life expectancy	Life expectancy indicating the number of years a newborn infant would live Source: World Bank Development Indicators
Patents applications	Number of patents applications Source: World Bank Development Indicators
Chinn-Ito index	A measure of capital account openness Source: http://web.pdx.edu/~ito/Chinn-Ito_website.htm
Investment rate	Investment as a share of GDP Source: World Bank Development Indicators
Investment growth rate	The growth rate of private investment Source: World Bank Development Indicators
Political constraints on the chief executive	Political constraints on the chief executive Source: Polity IV
Openness to trade	Import plus export, as a share of GDP Source: World Bank Development Indicators
Net equity inflows	Net inflows from equity securities including shares, stocks, depository receipts and direct purchases of shares in local stock markets by foreign investors Source: World Bank Development Indicators

Notes: Taiwan, China is not listed as a separate economy in World Bank Development Indicators. We compile data for Taiwan from other sources such as Penn World Table, Ministry of Education of Taiwan and Ministry of the Interior of Taiwan.

Table 3.3: Results of Replication and Extensions

	Dependent variable: real per capita GDP growth rate (%)					
	Only liberalizing countries				All countries that have data	
	Before-after difference (1)	Fixed effect (2)	Year effect (3)	Fixed and year effect (4)	OLS (5)	GMM (6)
Bekaert *et al*. (2005)	1.17***	1.24***	2.02***	1.05**	1.20***	0.97***
Replication	1.21***	1.50***	2.28***	1.08	1.19***	0.96***
All cases	1.43***	1.71***	1.52***	1.73**	1.12***	0.81***
All cases except special Eastern European Countries	1.13***	1.25***	1.17***	1.22*	0.89***	0.81***
All cases except China	1.40***	1.71***	2.02***	1.80**	1.01***	0.82***

Notes: Following Bakeart *et al*. (2005), two different samples are considered. One of them (the first four columns) only includes the countries that liberalized during 1980-1997. Column (1) reports the difference between the post-liberalization and pre-liberalization 5-year average growth rates. Column (2) through Column (4) report the estimates of the growth effects of equity market liberalization from three regressions with fixed effects, time effects and both, respectively. In each regression, the dependent variable is the annual growth rate of real per capita GDP and the independent variable is a dummy that takes a value of one if the stock market is liberalized and zero otherwise. There are no other control variables used. Column (5) and (6) represents the second sample, which depends on data availability: economic growth rates, the liberalization indicator as well as macroeconomic and demographic variables must all be available for the countries in this sample. It consists of both liberalizing countries and countries that never liberalized in the period under study. Column (5) reports the growth effect estimated from the OLS regressions. The dependent variable is non-overlapping five-year average growth rate. Three separate OLS regressions are carried out over three different periods (1981-1995, 1982-1996 and 1983-1997), and the simple average of the three coefficients are presented. Column (6) shows the results of a generalized method of moments (GMM). The dependent variable is overlapping five-year average growth rate. This GMM method allows researchers to better utilize the temporal components of the data and can adjust the standard errors for using overlapping observations. With different weighting matrices, it accommodates temporal heteroscedasticity, cross-country heteroscedasticity and/or SUR effects. *p<0.10, **p<0.05 ***p<0.01.

Table 3.4: Results of Cox Hazard Regressions

Dependent variable: Dummy variable indicating stock market liberalization			
	(1)	(2)	(3)
Real Per capita GDP ($1000)	-0.033**	-0.075**	0.139***
Government expenditure	-0.038	-0.076*	-0.188***
Investment rate	0.034**	0.011	0.025
Openness to trade	-0.015***	-0.018***	-0.020
Common law	1.936***	2.283***	3.158***
Inflation	0.0001	0.0007***	0.0007***
Chinn-Ito index		-0.157	-0.144
Private credit		0.075***	0.008
Constraints on executives		0.211***	0.158
Industry		0.071***	0.064*
Primary school enrollment		0.003	-0.002
Financial aid			2.362
Patent applications			-0.000
Observations	2753	2078	659
Liberalizations	40	36	22

Notes:

[a]This table reports the results of Cox proportional hazard regressions.

[b]Column (1), (2) and (3) refer to samples of 40, 36 and 22 countries, respectively.

[c]The dependent variable is a dummy indicating official liberalizations.

[d]All covariates are lagged by one year.

[e]*p<0.10, **p<0.05 ***p<0.01

Table 3.5: Covariates Balance Check

Covariate	Sample	Treated	Control	t-statistic	P-value
Real per capita GDP ($1000)	Unmatched	2971.1	1904.4	2.59**	0.01
	Matched	3023.6	2639.7	0.46	0.649
Government expenditure	Unmatched	12.691	14.29	1.59	0.113
	Matched	12.87	11.485	1.20	0.237
Investment rate	Unmatched	22.208	23.225	0.57	0.570
	Matched	22.546	26.948	1.62	0.113
Openness to trade	Unmatched	55.915	71.342	1.84*	0.066
	Matched	57.098	50.923	0.73	0.470
Common law	Unmatched	0.318	0.019	7.62***	0.000
	Matched	0.333	0.286	0.33	0.746
Chinn-Ito index	Unmatched	-0.678	-0.365	1.01	0.313
	Matched	-0.621	-0.400	0.55	0.582
Inflation	Unmatched	371.84	25.02	5.08***	0.000
	Matched	149.13	34.66	0.85	0.400
Private credit	Unmatched	37.32	24.29	3.00***	0.003
	Matched	38.51	31.58	0.97	0.339
Constraints on executives	Unmatched	5.36	2.20	1.16	0.246
	Matched	5.29	4.47	1.30	0.202
Industry	Unmatched	34.495	30.908	1.65*	0.099
	Matched	34.753	33.821	0.32	0.749
Primary school enrollment	Unmatched	102.68	99.37	0.94	0.349
	Matched	101.92	93.49	1.37	0.178
Financial aid	Unmatched	0.027	0.052	1.79*	0.075
	Matched	0.028	0.036	0.53	0.598
Patent applications	Unmatched	2815.4	542.7	5.00***	0.000
	Matched	2936.7	1219.1	1.22	0.229

Notes:

[a]The 13 variables are the same as in Col (3) of Table 3.

[b]The control group is generated using kernel (normal) matching.

[c]When Col (1) or Col (2) of Table 3 is used, covariate balance is easier to achieve due to a smaller number of covariates. Their balance check tables are available upon request.

[d]*p<0.10, **p<0.05 ***p<0.01

Table 3.6: Effects on Growth Using Matching

Results with 13 covariates N=22	Kernel (Normal)	Nearest Neighbor	Radius=0.1
1-year average	4.17**	4.36***	3.30*
2-year average	2.38*	2.69**	1.90*
3-year average	2.48*	2.15*	1.46
4-year average	2.23*	1.80*	1.16
5-year average	1.59	1.20	0.70
6-year average	1.52	0.71	0.57
7-year average	1.44	0.89	0.99
8-year average	1.21	0.89	0.86
9-year average	1.20	0.89	0.92
10-year average	1.38	1.11	0.95
Results with 11 covariates N=36	Kernel (Normal)	Nearest Neighbor	Radius=0.1
1-year average	2.50**	1.62	2.83***
2-year average	2.37***	2.64***	2.37***
3-year average	2.27***	2.63***	2.22***
4-year average	1.99***	2.10***	1.89***
5-year average	1.57***	1.52**	1.45***
6-year average	1.30***	1.16*	1.24**
7-year average	1.12**	0.89	1.12**
8-year average	1.03**	0.78	1.00**
9-year average	1.10**	1.23**	1.03**
10-year average	1.00**	1.22**	0.88**
Results with 6 covariates N=40	Kernel (Normal)	Nearest Neighbor	Radius=0.1
1-year average	2.25***	2.49**	2.32***
2-year average	1.92***	2.18**	1.92***
3-year average	1.74***	1.83**	1.72***
4-year average	1.45***	1.46**	1.45***
5-year average	1.00**	0.99*	1.01**
6-year average	0.91**	0.86	0.91**
7-year average	0.96**	0.89*	0.98**
8-year average	1.02**	1.00**	1.04**
9-year average	1.10***	1.14**	1.12***
10-year average	1.02***	1.19**	1.02***

Notes:

[a]This table represents results from propensity score matching.

[b]The data covers liberalization cases between 1980 and 1999. The first panel controls for the most covariates and has the fewest liberalization cases. This sample corresponds to column (3) of Table 3. The second panel corresponds to column (2) of Table 3. The third panel controls for the fewest covariates but has the largest sample which corresponds to column (1) of Table 3.

[c]As we seek to evaluate the average real GDP per capita growth effects over time for 10 years, countries 10 years or less before they liberalize cannot be included in the control group.

[d]The outcome variable is the change in the average growth rate before and after liberalizations.

[e]All covariates are lagged by one year to avoid simultaneous endogeneity.

[f]The post-liberalization period does not include the year of liberalization.

[g]The number of nearest neighbors is 2.

[h]We also experiment with other matching methods by using different numbers of nearest neighbors, applying different kernels and setting different radiuses. The results do not alter our major conclusions.

[i]*p<0.10, **p<0.05 ***p<0.01

Table 3.7: Effects on Growth Using Matching (All Cases)

Results with 13 covariates N=32	Kernel (Normal)	Nearest Neighbor	Radius=0.05
1-year average	2.42**	1.70	2.16
2-year average	1.67**	1.27	1.77*
3-year average	1.88***	1.70*	2.25***
4-year average	1.87***	2.05**	2.45***
5-year average	1.75***	1.76**	2.44***
6-year average	1.92***	1.92**	2.49***
7-year average	2.58***	2.64***	3.02***
8-year average	2.59***	2.63***	2.87***
9-year average	1.61**	1.59**	1.99**
10-year average	1.31**	1.45*	1.69**
Results with 11 covariates N=49	Kernel (Normal)	Nearest Neighbor	Radius=0.05
1-year average	1.85**	1.01	1.33
2-year average	1.18*	1.49	1.30*
3-year average	1.63**	2.24**	1.57**
4-year average	1.73***	2.59**	1.63**
5-year average	1.18**	1.34	1.01*
6-year average	1.21**	1.12	1.03*
7-year average	1.17**	1.16	1.19*
8-year average	1.13*	1.14	1.10*
9-year average	1.11**	1.08	1.12**
10-year average	1.12**	0.96	1.17**
Results with 6 covariates N=65	Kernel (Normal)	Nearest Neighbor	Radius=0.05
1-year average	0.90	1.03	0.94
2-year average	1.23**	1.71*	1.35**
3-year average	1.30*	1.71*	1.42**
4-year average	1.31**	1.38	1.41**
5-year average	1.22**	0.67	1.26**
6-year average	1.33**	0.88	1.35**
7-year average	1.38**	1.00	1.39**
8-year average	1.39**	0.97	1.40**
9-year average	1.18**	0.77	1.18**
10-year average	1.13**	0.76	1.11**

Notes:

[a]This table represents results from propensity score matching.

[b]The data covers liberalizations between 1980 and 2011.

[c]As we seek to evaluate the average real GDP per capita growth effects over time for 10 years, countries 10 years or less before they liberalize cannot be included in the control group.

[d]The outcome variable is the change in the average growth rate before and after liberalizations. e.g., the 5-year average growth rate change is 5-year post-liberalization average growth rate less 5-year pre-liberalization average growth rate.

[e]All covariates are lagged by one year to avoid simultaneous endogeneity.

[f]The post-liberalization period does not include the year of liberalization.

[g] ***,**,* denote significance at the 1, the 5, and the 10 percent level.

Table 3.8: Placebo Test

	Number of covariates=13	Number of covariates=11	Number of covariates=6
1-year average	-0.34	0.47	0.87
2-year average	-0.64	0.82	0.77
3-year average	-0.94	0.73	0.56
4-year average	-0.72	0.86	0.68
5-year average	-0.71	0.55	0.44
6-year average	-0.23	0.40	0.25
7-year average	0.06	0.29	0.15
8-year average	0.25	-0.06	-0.15
9-year average	0.22	-0.24	-0.33
10-year average	0.48	0.21	-0.00

Notes:

[a]This table represents the matching results as if the liberalizations had occurred ten years before the true liberalization dates.

[b]Kernel matching is applied, kernel being normal.

[c]*p<0.10, **p<0.05 ***p<0.01

Table 3.9: Results from Panel Regressions

Ctrls	(1)			(2)			(3)		
Regression	OLS	FE	Two-way FE	OLS	FE	Two-way FE	OLS	FE	Two-way FE
Lib0	0.13	0.19	0.16	0.14	0.38	0.46	0.11	0.70	0.83
Lib1	0.85	0.90	0.82	0.49	0.70	0.75	-0.04	0.25	0.40
Lib2	1.05**	1.11**	1.11**	1.13*	1.40**	1.57**	1.06	1.51	1.81*
Lib3	1.32***	1.42***	1.41**	0.82	1.23**	1.30**	1.27*	1.85***	2.25***
Lib4	0.55	0.68*	0.57	0.55	0.82	0.85	1.08**	1.49***	2.04***
Lib5	0.12	0.22	-0.11	-0.12	0.26	0.19	-0.41	0.28	0.47
Lib6	0.09	0.12	-0.17	0.04	0.42	0.32	-1.06	-0.49	0.01
Lib7	0.61	0.66	0.45	0.47	0.86	0.87	0.51	1.27*	1.98***
Lib8	0.47	0.52	0.11	0.29	0.56	0.25	0.46	1.01	1.25*
Lib9	0.91	0.89	0.43	1.23	1.47	1.09	-0.04	0.32	0.45
Lib10	-0.85	-0.83	-1.43**	-0.98	-0.74	-1.02*	-1.55	-1.09	-0.75
Obs	4276	4276	4276	3132	3132	3132	1118	1118	1118

Notes:

[a]Dependent variable: Real Growth Rate Per Capita

[b]Lib0 denotes the year of liberalization. Lib1 denotes 1 year after liberalization, so on and so forth

[c]White robust standard errors are used.

[d]Controls (1), (2) and (3) mean controlling for explanatory variables used in column (1), (2) and (3) of the table respectively.

[e]Data cover 1980-2010

[f] *p<0.10, **p<0.05 ***p<0.01

Table 3.10: Comparing Matching Results and Regression Results

Number of covariates=13	Kernel (Normal) Matching	Pooled OLS	FE	Two-way FE
1-year average	4.17	-0.04	0.25	0.40
2-year average	2.38	0.51	0.88	1.11
3-year average	2.48	0.76	1.20	1.49
4-year average	2.23	0.84	1.28	1.63
5-year average	1.59	0.59	1.08	1.39
6-year average	1.52	0.32	0.82	1.16
7-year average	1.44	0.34	0.88	1.28
8-year average	1.21	0.36	0.90	1.28
9-year average	1.20	0.31	0.83	1.18
10-year average	1.38	0.13	0.64	0.99
Number of covariates=11	Kernel (Normal) Matching			
1-year average	2.50	0.49	0.70	0.75
2-year average	2.37	0.81	1.05	1.16
3-year average	2.27	0.81	1.11	1.21
4-year average	1.99	0.75	1.04	1.12
5-year average	1.57	0.57	0.88	0.93
6-year average	1.30	0.49	0.81	0.83
7-year average	1.12	0.48	0.81	0.84
8-year average	1.03	0.46	0.78	0.76
9-year average	1.10	0.54	0.86	0.80
10-year average	1.00	0.39	0.70	0.62
Number of covariates=6	Kernel (Normal) Matching			
1-year average	2.25	0.85	0.90	0.82
2-year average	1.92	0.95	1.01	0.97
3-year average	1.74	1.07	1.14	1.11
4-year average	1.45	0.94	1.03	0.98
5-year average	1.00	0.78	0.87	0.76
6-year average	0.91	0.66	0.74	0.61
7-year average	0.96	0.66	0.73	0.58
8-year average	1.02	0.63	0.70	0.52
9-year average	1.10	0.66	0.72	0.51
10-year average	1.02	0.51	0.57	0.32

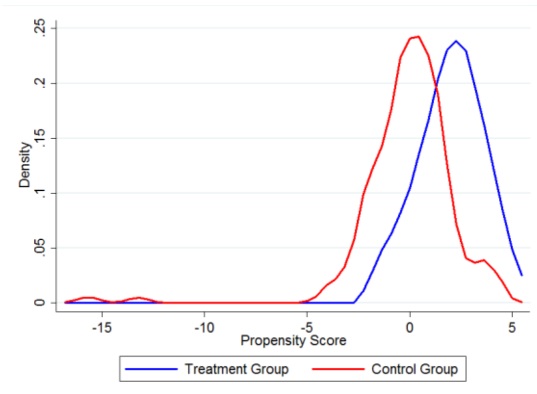
Notes:

[a]The second column is copied from the second column of Table 5.

[b]For the third, fourth and fifth column, the average growth effects are computed using the results in Table 6. Algorism: N-year average effect=$\frac{\sum_1^N \text{Libi}}{\text{N}}$, where Lib_i is the fitted value in each regression.

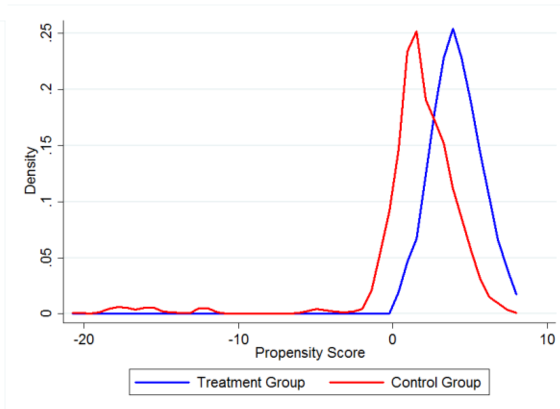
Figure 3.1: Distribution of Propensity Scores

Panel I



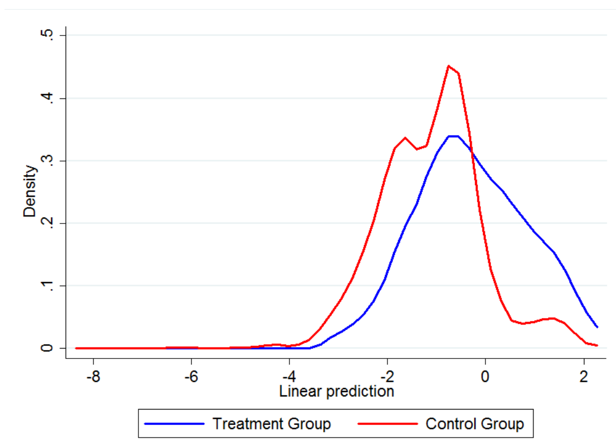
Kernel: Epanechnikov; No. of Treated Countries: 22.

Panel II



Kernel: Epan; No. of Treated Countries: 36.

Panel III



Kernel: Epan; No. of Treated Countries: 40.

References

Abadie, A. and J. Gardeazabal, "The economic costs of conflict: A case study of the Basque Country," *American economic review* 113-132, 2003.

Abadie, A., A. Diamond, and J. Hainmueller, "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*, 105, 493-505, 2010.

_________, "Comparative politics and the synthetic control method," *American Journal of Political Science*, 59, 495-510, 2015.

Anderson, Michael L., and Maximilian Auffhammer, "Pounds that kill: the external costs of vehicle weight," *The Review of Economic Studies,* 81: 535-571, 2014.

Andrade, E., F. Veloso, R. Madalozzo, and S. Ferreira, "Do borrowing constraints decrease intergenerational mobility? Evidence from Brazil," *Sao Paulo, Brazil: Banco Nacional de Desenvolvimento Econômico e Social (BNDES). Mimeographed document*, 2004.

Atkinson, A., A. Maynard, C. Trinder, J. Corlyon, S. Jenkins, and H. Sutherland, *Parents and children: incomes in two generations*, Heinemann Educational Books, London, 1983.

Azzam, H., *Emerging Arab Capital Markets*, Routledge, 2013.

Acemoglu, D., S. Johnson, and J. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 1369-1401, 2001.

Bae, K. and V. Goyal, "Equity market liberalization and corporate governance," *Journal of Corporate Finance*, 16, 609-621, 2010.

Banerjee, Abhijit V., and Esther Duflo, "The Economic Lives of the Poor." *Journal of Economic Perspectives,* 21, 141-168, 2007.

Barro, R., "Determinants of economic growth: a cross-country empirical study," NBER Working Paper No. w5698, 1996.

Baum, Christopher F, *An Introduction to Modern Econometrics Using Stata*. Stata Press, College Station, Texas, 2006.

Bekaert, G. and C. Harvey, "Capital flows and the behavior of emerging market equity returns," NBER Working Paper No. w6669, 1998.

Bekaert, Geert, and Campbell R. Harvey, "Foreign speculators and emerging equity markets," *Journal of Finance*, 55, 565-613, 2000.

Bekaert, G., C. Harvey, and C. Lundblad, "Emerging equity markets and economic development," *Journal of development Economics,* 66, 465-504, 2001.

Bekaert, G., C.Harvey, and C. Lundblad, “Equity market liberalization in emerging markets,” *Journal of Financial Research*, 26, 275-299, 2003.

Bekaert, G., C. Harvey, and C. Lundblad, “Does financial liberalization spur growth?” *Journal of Financial economics*, 77, 3-55, 2005.

Bekaert, G., C. Harvey, and C. Lundblad, “Growth volatility and financial liberalization,” *Journal of international money and finance*, 25, 370-403, 2006.

Ben Naceur, S., S. Ghazouani, and M. Omran, “Does stock market liberalization spur financial and economic development in the MENA region?” *Journal of Comparative Economics*, 36, 673-693, 2008.

Ben Gamra, S., “Does financial liberalization matter for emerging East Asian economies growth? Some new evidence.” *International Review of Economics & Finance*, 18, 392-403, 2009.

Bhaskar, V., "Sex selection and gender balance." *American Economic Journal: Microeconomics*, 3, 214-244, 2011.

Bjorklund, A. and M. Jantti, “Intergenerational income mobility in Sweden compared to the United States,” *American Economic Review*, 1009-1018, 1997.

Bjorklund, A., J. Roine, and D. Waldenström, “Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity?” *Journal of Public Economics*, 96, 474-484, 2012.

Bonfiglioli, A., and C. Mendicino, “Financial liberalization, banking crises and growth: assessing the links,” SSE/EFI working paper No. 567, 2004.

Bratberg, E., Ø. Nilsen, and K. Vaage, “Intergenerational earnings mobility in Norway: Levels and trends,” *Scandinavian Journal of Economics*, 107, 419-435, 2005.

Brown, Philip H., Erwin Bulte, and Xiaobo Zhang, "Positional spending and status seeking in rural China," *Journal of Development Economics* 96, 139-149, 2011.

Bulte, Erwin, Nico Heerink, and Xiaobo Zhang, "China's One-Child Policy and ‘the Mystery of Missing Women’: Ethnic Minorities and Male-Biased Sex Ratios," *Oxford Bulletin of Economics and Statistics* 73, 21-39, 2011.

Bumann, S., N. Hermes, and R. Lensink, “Financial liberalization and economic growth: A meta-analysis,” *Journal of International Money and Finance*, 33, 255-281, 2013.

Cajueiro, D., P. Gogas, and B. Tabak, “Does financial market liberalization increase the degree of market efficiency? The case of the Athens stock exchange,” *International Review of Financial Analysis*, 18, 50-57, 2009.

Case, Anne, Anu Garrib, Alicia Menendez, and Analia Olgiati. (2013). "Paying the piper: The high cost of funerals in South Africa." *Economic Development and Cultural Change* Vol. 62(1): 1-20.

Cavallo, E., S. Galiani, I. Noy and J. Pantano, "Catastrophic natural disasters and economic growth," *Review of Economics and Statistics* 95, 1549-1561, 2013.

Chai, Andreas and Wolfhard Kaus, “Signalling to whom? Conspicuous spending and the local density of the social group income distribution.” Working Paper No. 1218. *Papers on Economics and Evolution*, 2012.

Charles, Kerwin Kofi, Erik Hurst, and Nikolai Roussanov, "Conspicuous Consumption and Race," *The Quarterly Journal of Economics*, 124, 425-467, 2009.

Chen, Xi, Ravi Kanbur, and Xiaobo Zhang, "Peer effects, risk pooling, and status seeking: What explains gift spending escalation in rural China?" Working Paper, 2012.

Choe, H., B.Kho, and R. Stulz, “Do foreign investors destabilize stock markets? The Korean experience in 1997,” *Journal of Financial Economics*, 54, 227-264, 1999.

Corak, M. and A. Heisz, “The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data,” *Journal of Human Resources*, 504-533, 1999.

Couch, K. and T. Dunn, “Intergenerational correlations in labor market status: A comparison of the United States and Germany,” *Journal of Human Resources*, 210-232, 1997.

Cuong, N., “Which covariates should be controlled in propensity score matching? Evidence from a simulation study,” *Statistica Neerlandica*, 67, 169-180, 2013.

Das, M., and S. Mohapatra, “Income inequality: the aftermath of stock market liberalization in emerging markets,” *Journal of Empirical Finance,* 10, 217-248, 2003.

Dearden, L., S. Machin, and H. Reed, “Intergenerational mobility in Britain,” *Economic Journal*, 47-66, 1997.

Dehejia, R. and S. Wahba, “Propensity score-matching methods for nonexperimental causal studies,” *Review of Economics and statistics,* 84, 151-161, 2002.

Devereux, M. and G. Smith, “International risk sharing and economic growth,” *International Economic Review* 535-550, 1994.

Dube, A. and B. Zipperer, "Pooling Multiple Case Studies Using Synthetic Controls: An Application to Minimum Wage Policies," IZA Discussion Paper No. 8944, 2015.

Dunn, C., “The intergenerational transmission of lifetime earnings: Evidence from Brazil,” *BE Journal of Economic Analysis & Policy*, 7, 2007.

Eide, E. and M. Showalter, “Factors affecting the transmission of earnings across generations: A quantile regression approach,” *Journal of Human Resources*, 253-267, 1999.

Fan, Y., J. Yi, and J. Zhang, “Intergenerational mobility in China: Patterns and determinants,” working paper presented at Conference on the Study of Inequality in China, Beijing, 17 June, 2013.

Feng, Suwei, and Qiang Li, "Car Ownership Control in Chinese Mega Cities: Shanghai, Beijing and Guangzhou," *Journeys*, September, 40-49, 2013

Fuchs-Schündeln, N. and N. Funke, "Stock market liberalizations: Financial and macroeconomic implications," *Review of world economics*, 139, 730-761, 2003.

Galindo, A., F. Schiantarelli, and A. Weiss, “Does financial liberalization improve the allocation of investment?: Micro-evidence from developing countries,” *Journal of Development Economics*, 83, 562-587, 2007.

Gong, H., A. Leigh, and X. Meng, “Intergenerational income mobility in urban China,” *Review of Income and Wealth*, 58, 481-503, 2012.

Grawe, N., "Intergenerational mobility for whom? The experience of high-and low-earning sons in international perspective," in M. Corak (eds), *Generational income mobility in North America and Europe*, Cambridge University Press, Cambridge, 58-89, 2004.

Grier, R and K. Grier, "On the real effects of inflation and inflation uncertainty in Mexico," *Journal of Development Economics*, 80, 478-500, 2006.

Grier, K and R. Grier, “Only income diverges: A neoclassical anomaly.” *Journal of Development Economics*, 84, 25-45, 2007.

Grinblatt, Mark, Matti Keloharju, and Seppo Ikäheimo, "Social influence and consumption: Evidence from the automobile purchases of neighbors," *Review of Economics and Statistics*, 90, 735-753, 2008.

Guilmoto, Christophe Z, "Skewed sex ratios at birth and future marriage squeeze in China and India, 2005–2100," *Demography* 49, 77-100, 2012.

Guo, C. and W. Min, “Education and Intergenerational Income Mobility in Urban China,” *Frontiers of Education in China*, 3, 22-44, 2008 (In Chinese).

Gustafsson, B., “The degree and pattern of income immobility in Sweden,” *Review of Income and Wealth*, 40, 67-86, 1994.

Haider, S. and G. Solon, “Life-cycle variation in the association between current and lifetime earnings,” NBER Working Paper No. W11943, 2006.

Hassan, A., "Financial integration of stock markets in the Gulf: A multivariate cointegration analysis." *International Journal of Business*, 8, 335-346, 2003.

Heffetz, Ori, “A test of conspicuous consumption: Visibility and income elasticities,” *Review of Economics and Statistics*, 93, 1101-1117, 2011.

Henry, P., “Stock market liberalization, economic reform, and emerging market equity prices,” *Journal of Finance*, 55, 529-564, 2000a.

________, “Do stock market liberalizations cause investment booms?” *Journal of Financial economics*, 58, 301-334, 2000b.

Hermes, N., and R. Lensink, “Does financial liberalization influence saving, investment and economic growth? Evidence from 25 emerging market economies, 1973-96,” No. 2005/69. Research Paper, UNU-WIDER, United Nations University (UNU), 2005.

Hopkins, Ed, and Tatiana Kornienko, "Running to keep in the same place: consumer choice as a game of status," *American Economic Review*, 94, 1085-1107, 2004.

Hung, J., “Deregulation and liberalization of the Chinese stock market and the improvement of market efficiency,” *Quarterly Review of Economics and Finance*, 49, 843-857, 2009.

Iwata, S. and S. Wu, “Stock market liberalization and international risk sharing.” *Journal of International Financial Markets, Institutions and Money*, 19, 461-476, 2009.

Jaleel, F. and L. Samarakoon, “Stock market liberalization and return volatility: Evidence from the emerging market of Sri Lanka,” *Journal of Multinational Financial Management*, 19, 409-423, 2009.

Jantti, M. and E. Osterbacka, “How much of the variance in income can be attributed to family background? Evidence from Finland,” Unpublished working paper, 1999.

Jarvis, S. and S. Jenkins, “How much income mobility is there in Britain?” *Economic Journal*, 428-443, 1998.

Jayasuriya, S., “Stock market liberalization and volatility in the presence of favorable market characteristics and institutions,” *Emerging Markets Review*, 6, 170-191, 2005.

Kaus, Wolfhard, “Conspicuous consumption and “race”: Evidence from South Africa,” *Journal of Development Economics*, 100, 63-73, 2013.

Kawakatsu, H. and M. Morey, “Financial liberalization and stock market efficiency: an empirical examination of nine emerging market countries,” *Journal of Multinational Financial Management*, 9, 353-371, 1999.

Kaya, I., K. Lyubimov, and M. Miletkov, “To liberalize or not to liberalize: Political and economic determinants of financial liberalization,” *Emerging Markets Review*, 13, 78-99, 2012.

Kim, B. and L. Kenny, "Explaining when developing countries liberalize their financial equity markets," *Journal of International Financial Markets, Institutions and Money*, 17, 387-402, 2007.

Kouretas, G and M. Syllignakis, "Switching Volatility in Emerging Stock Markets and Financial Liberalization: Evidence from the new EU Member Countries," *Central European Journal of Economic Modelling and Econometrics*, 4, 65-93, 2012.

Lefranc, A. and A. Trannoy, “Intergenerational earnings mobility in France: Is France more mobile than the US?” *Annales d'Economie et de Statistique*, 57-77, 2005.

Levine, R., “Bank-based or market-based financial systems: which is better?” *Journal of financial intermediation*, 11, 398-428, 2002.

Lu, B., “Propensity Score Matching with Time Dependent Covariates,” *Biometrics*, 61, 721-728, 2005.

Manova, K., “Credit constraints, equity market liberalizations and international trade,” *Journal of International Economics*, 76, 33-47, 2008.

Mazumder, B., “Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data,” *Review of Economics and Statistics*, 87, 235-255, 2005.

Mitton, T., “Stock market liberalization and operating performance at the firm level,” *Journal of Financial Economics*, 81, 625-647, 2006.

Manova, K., "Credit constraints, equity market liberalizations and international trade," *Journal of International Economics*, 76, 33-47, 2008.

Mansur, A. and F. Delgado (eds), *Stock Market Developments in the Countries of the Gulf Cooperation Council*, Palgrave Macmillan, 2008.

Moav, Omer, and Zvika Neeman, "Status and poverty," *Journal of the European Economic Association*, 8, 413-420, 2010.

______________________________, "Saving Rates and Poverty: The Role of Conspicuous Consumption and Human Capital," *Economic Journal*, 122, 933-956, 2012.

O'Connor, T., "Equity market liberalization and firm growth," *Review of Development Finance*, 3, 1-12, 2013.

Osterbacka, E., "Family background and economic status in Finland," *Scandinavian Journal of Economics*, 103, 467-484, 2001.

Osterberg, T., "Intergenerational Income Mobility in Sweden: What do Tax-data Show?" *Review of Income and Wealth*, 46, 421-436, 2000.

Qian, Nancy, "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance," *The Quarterly Journal of Economics*, 123, 1251-1285, 2008.

Ranciere, R, A. Tornell, and F. Westermann, "Decomposing the effects of financial liberalization: Crises vs. growth," *Journal of Banking & Finance*, 30, 3331-3348, 2006.

Pascual, M., "Intergenerational income mobility: The transmission of socio-economic status in Spain," *Journal of Policy Modeling*, 31, 835-846, 2009.

Piraino, P., "Comparable estimates of intergenerational income mobility in Italy," *BE Journal of Economic Analysis & Policy*, 7, 2007.

Rosenbaum, P. and D. Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41-55, 1983.

Rosenbaum, P. and D. Rubin. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score," *The American Statistician*, 39, 33-38, 1985.

Rubin, D. and N. Thomas, "Matching using estimated propensity scores: relating theory to practice," *Biometrics*, 249-264, 1996.

Shi, X., J. Nuetah, and X. Xin, "Household income mobility in rural China: 1989–2006," *Economic Modelling*, 27, 1090-1096, 2010.

Solon, G., "Intergenerational income mobility in the United States," *American Economic Review*, 393-408, 1992.

________, "Cross-country differences in intergenerational earnings mobility," *Journal of Economic Perspectives*, 16, 59-66, 2002.

Stiglitz, J., "Capital market liberalization, economic growth, and instability," *World development*, 28, 1075-1086, 2000.

Ufier, A., "Quasi-experimental analytical on the effects of adoption of a value added tax," *Economic Inquiry*, 52, 1364-1379, 2014.

Wang, H., *Research on Income Distribution and Mobility in China*, Sun Yat-sen University Press, Guangzhou, 2007 (In Chinese).

Wei, Shang-Jin, and Xiaobo Zhang, "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," *Journal of Political Economy*, 119, 511-564, 2011.

Wei, Shang-Jin, Xiaobo Zhang, and Yin Liu, "Status competition and housing prices," *National Bureau of Economic Research Working Paper* No. w18000, 2012.

Yuan, Z. and C. Lin, "The Trend and Mechanism of Intergenerational Income Mobility in China: An Analysis from the Perspective of Human Capital, Social Capital and Wealth," *World Economy*, 36, 880-898, 2013.

Zimmerman, D., "Regression toward mediocrity in economic stature," *American Economic Review*, 409-429, 1992.