DISCUSSION PAPER SERIES

IZA DP No. 2524

Self-Selection and the Returns to Geographic Mobility: What Can Be Learned from the German Reunification "Experiment"

Anzelika Zaiceva

December 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Self-Selection and the Returns to Geographic Mobility: What Can Be Learned from the German Reunification "Experiment"

Anzelika Zaiceva
IZA Bonn

Discussion Paper No. 2524
December 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 2524
December 2006

ABSTRACT

Self-Selection and the Returns to Geographic Mobility: What Can Be Learned from the German Reunification "Experiment" *

This paper investigates the causal effect of geographic mobility on income. The returns to German East-West migration and commuting are estimated, exploiting the structure of centrally planned economies and a "natural experiment" of German reunification for identification. I find that the migration premium is insignificantly different from zero, the returns for commuters equal to 40 per cent, and the local average treatment effects for compliers are insignificant. In addition, estimation results suggest no positive self-selection on unobservables for migrants, and some evidence of positive self-selection on unobservables for commuters.

JEL Classification: F22, J61, R23

Keywords: returns to migration, causality, treatment effects

Corresponding author:

Anzelika Zaiceva
IZA
P.O. Box 7240
D-53072 Bonn
Germany
E-mail: zaiceva@iza.org

* I am grateful to Andrea Ichino for all his guidance, support and helpful advice, to Francis Vella for valuable comments and help with the programming and to Katharina Spiess for granting me access to the geo-code data. I also thank Holger Bonin, Hartmut Lehmann, Barry Chiswick, Jennifer Hunt, Joachim Frick, Nicola Fuchs-Schündeln, André Meier, Christopher Milde, David Jaeger and seminar participants at IZA, European University Institute, IZA Summer School, Second IZA Migration Meeting, IV BRUCCHI LUCHINO Labour economics workshop, ESPE meeting in Verona, EEA meeting in Vienna and EALE meeting in Prague for their comments. All remaining errors are mine. An earlier version of this paper was published as DIW Discussion Paper No. 580.

1 Introduction

With cumulative net migration of 7.5 per cent of the original population over the period 1989-2001, East Germany has the second highest emigration rate (after Albania) among the countries formerly behind the Iron Curtain (Brücker and Trübswetter, 2004, Heiland, 2004). The emigration rates have tended to increase again since 1997, and there seems to be no sign of income convergence from 1995 onwards (Figure 1 and OECD, 2001). Moreover, due to the particular geography of Germany, commuting to the West is a popular option for those who do not want to incur fixed costs of migrating, and it may substitute for emigration. Since geographic mobility constitutes an investment in human capital, these phenomena have raised concerns that individuals with high abilities move to the West ("brain drain"), as well as raising the question of how large the mobility premium is in the West. Such issues are also gaining general importance in light of the eastern enlargement of the European Union in May 2004 and resulting European East-West migration.

In an attempt to answer these questions, it is important, however, to separate the pure effect of geographic mobility from the effect of confounding factors. The reason why doing this is difficult is often attributable to the unavailability of the relevant data and credible exclusion restrictions.

This paper attempts to fill the gap and to estimate the *causal* effect of geographic mobility on income. In its main contribution to the literature it exploits the structure of the centrally planned economy of the former German Democratic Republic (GDR) together with the unique event of German reunification in order to make causal statements about the returns to geographic mobility from East to West Germany, controlling for the potential self-selection on unobservables.

Migration theory (Roy, 1951, Borjas, 1987) postulates that migrants will be positively selected if the distribution of earnings is more unequal in the destination region than in the origin.[1] There exists a vast empirical literature on migration, in which authors have investigated the selectivity issue, using standard Heckman's procedure, or have documented the association between migration and income. The majority of the existing empirical studies on East-West German migration address the question of self-selection indirectly.[2] The first study that explicitly deals with this issue is a recent paper by Brücker and Trübswetter (2004), in which the authors find no robust evidence of positive self-selection on unobservables for migrants from 1994 to 1997. As for the mobility premium, Hunt (2001) shows that those who took a job in the West between 1990 and 1991 enjoyed large wage gains, but that the correlation between wage growth and working in West Germany is small and insignificant for the subsequent movers. She notes that an economy undergoing a successful transition would initially have high returns to moving, which would fall as the transition progressed.

This paper exploits programme evaluation techniques and attempts to identify the effect of treatment (geographic mobility) on the treated (mover) as well as the so-called local average treatment effect for compliers (a subpopulation of movers whose status

[1] Chiswick (1999) shows that Roy's model is a special case of the human capital model of migration (Sjaastad, 1962).

[2] Burda (1993), Burda et al (1998) analyze individuals' intentions to move West. Hunt (2006) estimates the reduced form multinomial logit of the decisions to move, to commute or to stay.

changes with the instrument). I investigate these questions using both parametric and nonparametric econometric methodologies.

Home ownership and geographic residence before unification are argued to provide the exogenous sources of variation in migration and commuting, respectively, since housing decisions and voluntary geographic labor mobility were usually restricted in the former GDR. Under communism, an elaborate plan directed the allocation of inputs, the distribution of outputs, and wage levels. Usually rules and party membership played an important role. Moreover, German reunification was not anticipated by anybody until shortly before the event. Although one may still argue that the allocation of housing and residence of individuals in the communist economy was not random, it was largely based on factors that are not relevant for the market economy and post-unification individual incomes, which are thus ignorable.

The main findings of this paper are as follows. First, no evidence of positive selection on unobservables for migrants and some evidence of positive self-selection for commuters is found. Second, the returns in terms of long-run income are insignificant for both migrants and compliers. The returns for commuters are high and equal to 40 per cent, but the local average treatment effect for compliers is insignificant.

The paper is organized as follows. Section 2 provides the description of the data and section 3 justifies the instruments. Section 4 outlines the estimation strategy. Estimation results are discussed in section 5, and section 6 provides a sensitivity analysis. Section 7 concludes.

2 Data, Definitions and Sample Selection

The data used in this paper are extracted from the public use file of the representative German panel household survey (GSOEP)[3] and merged with the confidential geographical coding of individual places of residence. Due to the GSOEP's longitudinal structure it is possible to identify and trace movers (and their incomes). Another advantage of this dataset is that the first wave of the eastern sample was drawn in June 1990, i.e. before the monetary union and formal unification took place, and thus provides a unique opportunity to use pre-unification data to construct the exogenous source of variation in mobility. The main disadvantage of the dataset is the small number of observations for movers.

An individual is defined as a migrant if he has changed his residence from East to West Germany at least once during 1990-2001; otherwise he is a stayer. An individual is a commuter if he lives in the East and his region of work is West Germany in any of the years 1990-2001.[4] A definition of income is not trivial in such a study. Theory suggests that while making a decision to move, an individual takes his total lifetime income into account, and empirical studies find that the assimilation period matters.[5] In order to be consistent with the theoretical definition of lifetime "permanent" income, as well as wanting to avoid the problem of a transitory income drop right after the move, and to save observations, I have used the mean of annual incomes as a dependent variable. I thus average over

[3] See SOEP Group (2001).

[4] Note that when defining migrants in this way I have to include commuters within "stayers", and when defining commuters - actual and potential migrants within "stayers". I also experiment with excluding actual and potential movers from the respective comparison groups (section 6).

[5] It is argued that estimates based on earnings data with limited time horizons will not capture life-cycle wage growth, tending to downward bias in the estimated returns (Greenwood, 1997).

the available years for stayers, over the available years after an individual migrates for migrants, and over the years during which an individual commutes for commuters. The total annual income is defined as a sum of labor income (wages, second-job and self-employment earnings) and various social security benefits (such as unemployment benefits, maternity benefits etc.). The mean income is only set to missing if information on all the components is missing.[6] All incomes are inflated to 2001 by regional CPIs and are expressed in DM.

I restrict the sample to easterners who were living in East Germany in 1990, exclude pensioners and students, and use the incomes of individuals who are at least 18 years old in each year.[7] Final sample sizes in the most restricted specifications are 3,043 observations for migration (of whom around 6 per cent are migrants[8]), and 2,953 observations for commuting (of whom around 15 per cent are commuters).

The instruments used in this study are as follows. For the migration equation, I construct a dummy which equals one if an individual was a home-owner in 1990, and is zero otherwise.[9] For the commuting equation, the instrument is proximity to the former East-West German border and equals one if an individual resided in a county ("*kreis*") that was on this border in 1990.[10] Both instruments approximate theoretical

[6] I also exclude the obvious outliers from the sample, i.e. individuals whose average annual income is less than 1,000 DM (19 observations) or greater than 130,000 DM (5 observations). I have experimented with different thresholds and kept all individuals in the sample, and the results were not affected.

[7] I also drop the return and multiple migrants here (around 20 per cent), but retain them in the robustness checks (section 6).

[8] This number is consistent with the aggregate figures.

[9] 32 per cent of the respondents in 1990 in East Germany reported owning a house / flat.

[10] Around 30 per cent of East Germans lived in such "border counties" (including Berlin) in 1990. I have experimented with different definitions of proximity, including counties with a common border, additional counties within 50 and 30 kms. from the border and other, and have selected this one because it generated the strongest instrument. I have also dropped Berlin from the sample in the robustness

costs of moving: the former captures the well-established negative relation between home ownership and the propensity to migrate, while the later captures the costs of commuting West that increase with distance from the border.

Kernel densities of average total annual incomes for movers and stayers are shown in Figure 2. As expected, the distribution of incomes for stayers is more compressed, and there are more migrants and commuters in the upper tail of income distribution. Descriptive statistics for the key variables is given in Table 1. All potential movers have on average a higher total annual income than stayers. Compared to stayers, migrants tend not to own a house in 1990, and commuters tend to live in the border regions in 1990. As expected, potential movers are younger, single and better educated than stayers. There are more males among commuters, however, more females among migrants. Table 1 presents some systematic differences in observable characteristics between movers and stayers; thus, there is a reason to suspect, a priori, that selection on unobservables will be an issue. To cope with this, I rely in the remainder of the paper on the instrumental variables, which are justified in the next section.

3 Are the Instruments Legitimate?

In order to make causal statements about the returns to geographic mobility, it is important to justify the validity of the instruments. Unfortunately, this assumption cannot be tested, and one has to rely on the available general facts. To be a valid instrument,

checks.

pre-unification home ownership and residence dummies must affect income only through migration or commuting, i.e. they must be uncorrelated with any non-ignorable confounding factors that affect ex-post income in the market economy, such as ability or motivation. This can be justified by referring to the structure of centrally planned economies.

In the GDR, as in any communist society, there was a high degree of centralization in the labor and product markets: all firms were owned by the state and an elaborate plan directed the allocation of inputs, the distribution of outputs, wage levels and prices (Krueger and Pischke, 1995). To secure constant prices for inhabitants, the state bore 80 per cent of costs of basic supplies, from bread to housing. Shortages were the norm. The distribution of income was compressed, and wage inequality was very low.[11] Official unemployment was absent, since workers were kept inefficiently in companies even if they were unproductive. Political tolerance was important: the system only functioned smoothly when its component parts were staffed with individuals whose values coincided with those of the regime. In general, the communist ideology stressed uniformity of outcomes, irrespective of individual differences in ability or effort.

Housing and occupational choices, and thus voluntary geographic labor mobility, were restricted. In principle, everyone had a right to a house; however, due to rationing by the state (the so-called System of Material Balances), long queues were the norm.[12] Access to housing was regulated largely through informal (and often politically mediated) networks.

[11] Fuchs-Schündeln and Schündeln (2005) report that in 1988, the average net income of individuals with a university degree was only 15 per cent higher than that of blue-collar workers.

[12] The "waiting list" could be very long. For example, the wait for an apartment in the Soviet Union during the 1980s was typically 10 to 15 years; as a result, families had to plan and buy housing for their children to live in in advance.

In many ways access to material and social activities in the ex-GDR was mediated through the sphere of work, and, in particular, the FDGB unions acted as the prime political link between the working population and the Socialist power elite, and as key agents in the distribution of housing. In general, flats were allocated to individuals due to urgent need or merit, personal connections or corruption, or by inheritance. Those who paid a nominal rent for a state-owned flat enjoyed considerable consumer surplus (Kornai, 1980). As for the occupational choice, job offers were usually made to individuals immediately after completion of their education and according to the Socialist plan. Even admissions to the various fields were regulated by the plan.[13]

Overall, the communist system operated like a large internal labor market, with rules and party membership playing an important role in the allocation of jobs and wages (Krueger and Pischke, 1995). As a result, little was left to individual abilities and motivation. Finally, the fall of the Berlin Wall in 1989 could not been foreseen. Therefore, to the extent that individuals had not been self-selecting into home ownership status or into the regions on the basis of their unobservable characteristics relevant for the market economy, the instruments provide the exogenous source of variation in mobility, and the assignment to treatment is strongly ignorable.

However, the exclusion restriction assumption is violated if, for example, more able persons were also more successful in gaining access to their own housing, leading to an upward bias in the estimates. Moreover, in the former GDR, only those who supported

[13] Only a certain quota of students was allowed to complete the last two years of high school, necessary to attend university. Additional criteria were membership in the official youth organization, political tolerance, and family background (Fuchs-Schündeln and Schündeln, 2005).

the regime (i.e. party members and the so-called "nomenklatura") were likely to be allowed to live close to the western border. If these people also were more motivated, the validity of the instrument will be violated unless one controls for the "nomenklatura effect". Fortunately, Bird, Frick and Wagner (1998) provide a proxy for party membership and nomenklatura status - telephone availability before unification, which I also use in the robustness checks (see section 6).

Finally, an informal exercise can be undertaken to further justify the instruments. If they approximate a randomized experiment, the characteristics of those for whom the instrument equals one must be equal to those for whom it equals zero, meaning that persons are randomly assigned across the two groups. Table 2 shows[14] that for migration the home ownership dummy is indeed orthogonal to some covariates, yet there exist differences (at 5 per cent) in some of them. Contrary to expectations, however, the more educated and those having a higher pre-treatment income are *less* likely to own a house before unification. Thus, it is likely that housing was not randomly allocated to individuals in the communist economy, however such allocation was probably based on some political factors and personal connections (or corruption) and not on the unobservables that are relevant for the market economy, such as individual ability and motivation. Moreover, differences in most characteristics, although statistically significant, are not economically pronounced[15]. For commuting, the border dummy is orthogonal to all covariates with the exception of telephone availability in 1990, which actually confirms the existence of the "nomenklatura effect".

[14] Note, however, that these are only observable characteristics.

[15] Differences in all characteristics range from 9 to 20 per cent of the respective standard deviations.

Therefore, although one may still argue that the allocation of housing and residence of individuals in the Communist economy was not random, it was largely based on factors that are not relevant to the market economy and post-unification individual incomes. Overall, I believe that the evidence presented in this section allows for a valid causal inference, at least for commuters.

4 Econometric Methodology

In order to estimate the causal effect of geographic mobility on the income of movers, a standard potential outcomes model is used. Let $D_i = 1$ if individual is a mover, and $D_i = 0$ otherwise. Then the average effect of treatment on the treated (ATT) can be written as follows:

$$\begin{aligned} ATT &= E(\Delta_i | Z_i, D_i = 1) = E(Y_{1i} - Y_{0i} | Z_i, D_i = 1) = \\ &= E(Y_{1i} | Z_i, D_i = 1) - E(Y_{0i} | Z_i, D_i = 1) = \\ &= E(\Delta_i) + E(\eta_i | Z_i, D_i = 1) \end{aligned} \tag{1}$$

where Z_i are individual socio-economic characteristics and exogenous variables, Y_{1i} and Y_{0i} are individual i's potential incomes with and without movement.

ATT is the difference between actual outcome for movers and a counterfactual outcome for movers had they stayed. It equals to the average effect for a random person in the population *plus* the idiosyncratic gain from treatment (the returns to unobservables), and there is no a priori reason to expect $E(\eta_i|Z_i, D_i = 1) = 0$. Thus, the OLS estimation of (1) provides biased and inconsistent estimates.

To calculate the effect of moving West on income, I first estimate a parametric sample selection model of Heckman (1976, 1979). Note that this procedure requires exclusion restrictions. In addition, if the joint normality assumption does not hold, it produces inconsistent estimates. Then, I also estimate the nonparametric sample selection model of Das, Newey and Vella (2003) that does not impose any distributional assumptions and does not restrict the form of the correction function. The identification requires exclusion restrictions, and the model is identified up to an additive constant. The approach amounts to estimating in the first step a conditional probability of selection (propensity score) without making any distributional assumptions, and, in the second step, to approximating the correction function with polynomial series. The order of the correction term is chosen using a leave-one-out cross-validation criterion. I also use two semiparametric techniques to consistently estimate the intercept (Heckman, 1990 and Andrews and Schafgans, 1998). The ATT is then calculated as the difference between the actual outcome for movers and the counterfactual outcome for movers had they stayed. Finally, making no restrictions on unobserved heterogeneity, I also estimate the local average treatment effect (LATE) for compliers (Angrist, Imbens and Rubin, 1996).[16]

[16] Note that the Random Assignment, Exclusion Restrictions and a Non-zero Effect of the Instrument on the Treatment assumptions are satisfied based on the evidence presented in Sections 3 and 5. SUTVA assumption seems plausible, since movers constitute only a small fraction of the population, thus ruling out general equilibrium effects. Finally, the assumption of Monotonicity (no defiers) also seems plausible, since both owning a house and living far from the border constitute costs for mobility.

5 Estimating the Effect of Mobility on Income

I use the standard semi-log specification of the income function. Variables such as experience, education and marital status in 2001 are endogenous; thus, only exogenous variables, such as sex, age and its square (as a proxy for experience), the predetermined marital status (as a proxy for "psychic" migration costs) and human capital variables in 1990 are used.

5.1 Returns to Migration

The first stage estimates (available upon request) confirm that, on average, home owners are less likely to migrate and that the instrument is strong (see also Table 5). Probit marginal effects indicate that the probability of moving West decreases with age, males are less likely to migrate, and both university degree and marital status have expected signs, but neither these variables nor occupation variables, nor the state‘s unemployment rate are significant. Heckman‘s second stage estimates (Table 3) suggest that males have a higher total income than females, experience as proxied by age and its square has the traditional concave profile, and university graduates earn more. However, neither vocational education nor occupational dummies are significant for movers, suggesting that part of the human capital acquired in the centrally planned economy is not transferable to the West. The coefficient on the inverse Mills ratio for movers is positive, but insignificant, indicating no evidence of significant positive self-selection after having controlled for human capital and demographics. Estimates for stayers suggest that, on average, male stayers have a higher total income than females, university graduates earn more, experience has the expected sign, those who had a vocational degree and were employed in the government sector in 1990 earn more, and those in blue-collar occupations in 1990 earn less. The Mills ratio for stayers is also insignificant.

To test the normality assumption I use the conditional moment test (Newey, 1985, Pagan and Vella, 1989), which indicates that normality cannot be rejected, implying that Heckman‘s estimates are consistent. Nevertheless, I also experiment with the nonparametric sample selection model and do not restrict the form of the correction function. In the first stage, I estimate a linear probability model and construct predicted probabilities. The cross-validation criterion suggested the linear correction function for movers and a polynomial of order 3 for stayers. Table 3 also shows the nonparametric second stage estimates. The coefficients on covariates for both stayers and movers are similar to the parametric ones. When normality is not imposed, there is again no evidence of positive self-selection for movers.

Finally, I estimate the model by IV-LATE framework. Table 5 (panel A) summarizes the so-called intention-to-treat effects (reduced form migration and income equations, columns 1-2), and structural IV estimates (column 3). The IV point estimate is not statistically significant. The local average treatment effect for compliers shows that those individuals who migrated if they did not own a house in 1990, but would not have migrated if they had owned a house, have no significant returns to their ex-post long-run income from migration.

Table 6 (panel A) summarizes treatment effects for migrants in the different econometric models used.[17] The effects of migration for both migrants and compliers are not statistically different from zero. One should bear in mind, however, that the results for migration have to be interpreted with caution: there might still exist some doubts about the validity of the instrument, the standard errors in IV are generally very large and the coefficients flip from large negative to positive.

Overall, several interesting findings occur from the estimates. First, no evidence of positive self-selection on unobservables for East-West German migrants during 1990-2001 is found. Such a result is partly in line with Brücker and Trübswetter (2004), and is also consistent with the theoretical predictions of the human capital model (Chiswick, 1999), when direct out-of-pocket costs of migration are small. Given that the inequality of earnings in East Germany has approached West German levels in the late 1990s, the standard Roy‘s model would also predict that a positive selection bias should disappear.[18]

Second, both treatment effect for migrants and the LATE for compliers are insignificant. This result might be a consequence of high unemployment in the East when people move West not in search of a higher income but to escape from unemployment, and it may also be the cause of return migration to the East. Together with no positive selection for migrants it may also reflect attitudes towards risk or non-transferable human capital. Finally, the exclusion of earlier migration (1989-1990) from the analysis due to the unavailability of data may bias the effects downward, since high initial migration most

[17]Standard errors of the effects for sample selection models are calculated as for the Oaxaca decomposition.

[18]Ideally, however, one should estimate year by year regressions in order to document the evolution of the selection bias over years, since the cohort quality effect might be at work here, the first migrants being of better quality than the subsequent movers. Unfortunately, small number of observations prewent me from doing this.

probably left behind those with the highest migration costs. These results, however, are not entirely surprising. Hunt (2001), for instance, also finds that the correlation between wage growth and working in the West is insignificant for the post-1991 migrants.

5.2 Returns to Commuting

Reduced form estimates for commuters (available upon request) suggest that on average males, the young and university graduates are more likely to commute West. The West border dummy has a large positive impact on the probability of commuting (i.e. the instrument is strong, see Table 5) and indicates that the costs of commuting indeed increase with the distance. Second-stage Heckman's estimates (Table 4) suggest that males and university graduates earn more, and experience has a traditional profile. For stayers, in addition, being employed in the government sector and having a vocational degree in 1990 affect their ex-post incomes positively, while being a blue-collar employee in 1990 affects it negatively. The selection correction terms are insignificant for both commuters and stayers. However, the conditional moment test rejects the normality assumption, implying that parametric estimates are inconsistent.

In the nonparametric model, the leave-one-out cross-validation criterion suggested a polynomial of order 2 for commuters and no correction polynomial for stayers. The estimated coefficients for both commuters and stayers are again similar to those in the parametric model, apart from the correction terms. In addition, the marginal effects of the correction functions for commuters are positive, thus suggesting positive self-selection for commuters.

Panel B of Table 5 shows the intentions-to-treat effects and IV estimates. Again, IV point estimates are not statistically significant. Hence, the local average treatment effect for individuals who commute if they were living in the border regions in 1990 and who would not have commuted otherwise, is not statistically different from zero.

Table 7 (panel A) summarizes all the effects.[19] Overall, for commuters, positive self-selection seems to be present. The LATE for compliers is again insignificant. However, the treatment effect for commuters equals 0.4, suggesting a large 42 per cent effect on the average long-run income.

6 Robustness Checks

The following sensitivity analysis was undertaken. First, I check how robust the results are to the inclusion of additional controls. I include a dummy which equals one if a person was unemployed in 1990 to check how the lagged employment status influences both the decision to move and ex-post incomes. I then add the household monthly income in 1990 in order to capture additional household-level characteristics. Second, I improve on the validity of the instruments controlling for the "nomenklatura" effect mentioned above. One may argue that it is also important to control for the ideology, thus I also include a variable that ranks political interests of a person before unification. Finally, I control for the lagged hours worked per week. Third, I exclude the self-employed from the sample, since there might be self-selection into this group. Fourth, I retain all return

[19] Standard errors of the effects for sample selection models are calculated as for the Oaxaca decomposition.

and multiple movers in the sample. Fifth, I improve the definition of the control group: I drop commuters from the control group for migrants, and migrants from the control group for commuters. Finally, I also control for the years from which the income is taken in order to take further account of wage convergence. Panel B of Tables 6 and 7 shows these sensitivity checks for migration and commuting equations, respectively. In general, the effects are similar to those reported in Panel A, and are more robust for commuting equation.[20]

One could still argue that the income *growth* and not income per se is a relevant dependent variable as it differences away any fixed effects in income levels. However, it still leaves the selection bias associated with the non-random selection of movers, thus it is still necessary to rely on valid exclusion restrictions in order to get rid of the bias. In panel C of Tables 6 and 7 all models have been reestimated using income growth as a dependent variable.[21] The results have not changed much. The resulting treatment effects for migrants are again insignificant across all the models. For commuters, a consistent nonparametric model suggests ATT equal to 29 per cent.

[20] In addition, all models have been re-estimated without human capital covariates and have generated qualitatively identical results (available upon request). Also, I have used labor income as a dependent variable. The results for migration were qualitatively the same. For commuters, nonparametric estimates were slightly higher (0.46) and LATE for compliers was marginally significant and equal to 0.4. These results seem to suggest that commuting particularly pays off with respect to the labor income, which is, in fact, true by definition of commuters. Finally, I have reestimated the models for two periods, 1990-1995 and 1996-2001, as well as excluding Berlin from the sample. All results are available upon request.

[21] The growth variable is constructed as follows. First, for migrants, I average over the available years before and after an individual move, and for commuters - over the years before the first commuting and after it, and construct $income_i^b$ and $income_i^a$, respectively. I then identify the so-called "average" year weighted by the number of individuals who move before and after it. Then I average the incomes before and after that year for stayers. Finally, I construct $income\ growth_i = \ln(income_i^a) - \ln(income_i^b)$.

7 Conclusions

The question of the returns to geographic mobility, especially in the context of transition economies, remains difficult to deal with, mainly due to data availability and identification problems. This chapter exploited a structure of the centrally planned economy of the ex-GDR and a "natural experiment" of German reunification, and attempted to make a causal inference for the returns to East-West German migration and commuting. Preunification home ownership was argued to provide an exogenous source of variation in migration, and proximity to the West German border before unification in commuting.

The main findings are as follows. First, no evidence of positive selection on unobservables for migrants and positive self-selection for commuters was found. Second, no significant returns to migration in terms of long-run income seem to exist. One should bear in mind, however, that the findings for migration have to be interpreted with caution. The returns for commuters are high and equal approximately 40 per cent, however, they are also insignificant for compliers. A higher overall gain for commuters is in line with expectations, taking into account the higher costs of migration and lower unemployment rate for commuters than for migrants. This may also suggest that commuting might indeed be a substitute for migration. Third, the results (especially for commuters) are robust to different changes in specifications and in the sample.

Overall, migrating West does not appear to be a significantly rewarding option for eastern Germans in the long run. This fact, although subject to the assumptions and definitions used in this study, could constitute an important part of the explanation of East-West migration in Germany.

References

[1] Andrews, D.W., and M.A. Schafgans, 1998, Semiparametric estimation of the intercept of a sample selection model, Review of Economic Studies 65, 497-517.

[2] Angrist, J.D., G. W. Imbens, and D.B. Rubin, 1996, Identification of causal effects using instrumental variables, Journal of the American Statistical Association 91, 444-472.

[3] Bird, E.J., J.R. Frick, and G.G. Wagner, 1998, The income of socialist upper class during the transition to capitalism: Evidence from longitudinal East German data, Journal of Comparative Economics 26, 211-225.

[4] Borjas, G.J., 1987, Self-selection and the earnings of immigrants, American Economic Review 77, 531-553.

[5] Brücker, H., and P. Trübswetter, 2004, Do the best go West? An analysis of the self-selection of employed East-West migrants in Germany, Discussion Paper no. 986, Instituts zur Zukunft der Arbeit (IZA), Bonn, Germany.

[6] Burda, M.C., 1993, The determinants of East-West migration: Some first results, European Economic Review 37, 452-461.

[7] Burda, M.C., and J. Hunt, 2001, From reunification to economic integration: Productivity and labor markets in Germany, Brookings Papers on Economic Activity 2, 1-71.

[8] Burda, M.C., W. Härdle, M. Müller, and A. Werwatz, 1998, Semiparametric analysis of German East-West migration intentions: Facts and theory, Journal of Applied Econometrics 13, 525-541.

[9] Chiswick, B.R., 1999, Are immigrants favorably self-selected?, American Economic Review 89, 181-185.

[10] Das, M., W.K. Newey, and F.G. Vella, 2003, Nonparametric estimation of sample selection models, Review of Economic Studies 70, 33-58.

[11] Fuchs-Schündeln, N., and M. Schündeln, 2005, Precautionary savings and self-selection - Evidence from the German reunification "experiment", Quarterly Journal of Economics 120, 1085-1120.

[12] Greenwood, M.J., 1997, Internal migration in developed countries, in: M.R. Rosenzweig and O. Stark, eds., Handbook of population and family economics, Vol. 1b (North-Holland, Amsterdam).

[13] Harris, J.R. and Todaro, M.P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. The American Economic Review 60: 126-42.

[14] Heckman, J.J., 1976, The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, Annals of Economic and Social Measurement 15, 475-492.

[15] Heckman, J.J., 1979, Sample selection bias as a specification error, Econometrica 47, 153-162.

[16] Heckman, J.J., 1990, Varieties of selection bias, American Economic Review 80, 313-318.

[17] Heiland, F., 2004, Trends in East-West German migration from 1989 to 2002, Demographic Research 11, 173-194.

[18] Hunt, J., 2001, Post-unification wage growth in East Germany, Review of Economics and Statistics 83,190-195.

[19] Hunt, J., 2006, Staunching emigration from East Germany: Age and the determinants of migration, Journal of the European Economic Association (forthcoming).

[20] Kornai, J., 1980, Economics of shortage, (North-Holland, Amsterdam).

[21] Krueger, A.B., and J-S. Pischke, 1995, A comparison of East and West German labor markets before and after unification, in: R.B. Freeman and L.F. Katz, eds., Differences and changes in wage structures, (University of Chicago Press, Chicago).

[22] Newey, W.K., 1985, Maximum likelihood specification testing and conditional moment tests, Econometrica 53, 1047-1070.

[23] OECD, 2001, Economic surveys Germany, (Organisation for Economic Co-operation and Development Press, Paris).

[24] Pagan, A., and F.G. Vella, 1989, Diagnostic tests for models based on individual data: A survey, Journal of Applied Econometrics 4, suppl. S29-S59.

[25] Roy, A.D., 1951, Some thoughts on the distribution of earnings, Oxford Economic Papers 3, 135-146.

[26] Sjaastad, L., 1962, The costs and returns of human migration, Journal of Political Economy 70, No.5, Part 2, 80-93.

[27] SOEP Group, 2001, The German socio-economic panel (SOEP) after more than 15 years - Overview, in: E. Holst, D.R. Lillard and T.A. DiPrete, eds., Proceedings of the 2000 fourth international conference of German socio-economic panel study users (GSOEP2000), Vierteljahrshefte zur Wirtschaftsforschung, Jg 70, Nr. 1, S. 7-14.

Table 1: Descriptive statistics

	Migration		Commuting	
	Migrants	Stayers	Commuters	Stayers
average annual income	39754	31125	43128	30009
	(26828)	(16937)	(22084)	(16739)
home owner, 1990	0.16	0.33		
border with West, 1990			0.48	0.27
gender	0.42	0.52	0.65	0.49
age	26.08	31.93	28.59	32.05
	(11.36)	(11.53)	(11.07)	(11.67)
spouse	0.61	0.74	0.69	0.74
university degree	0.16	0.09	0.13	0.09
vocational education	0.78	0.88	0.83	0.88
government sector	0.44	0.33	0.31	0.34
blue-collar employee	0.26	0.35	0.40	0.33
telephone	0.23	0.23	0.28	0.22
state's unempl. rate, 1992	10.51	10.49	10.70	10.45
	(1.02)	(0.93)	(1.02)	(0.91)

Note: standard deviations in parentheses. All time-variant demographic and human capital variables are of 1990. Incomes are annual, inflated by regional CPIs to 2001 and expressed in DM. Minimum sample sizes are 3043 observations for migration, and 2953 observations for commuting. Average annual income is a sum of labor income (wages, second job and self-employment income) and social security benefits (such as unemployment benefits, maternity benefits etc).

Table 2: Means of the variables by instruments

	Migration		Commuting	
	home owner in 1990	not home owner in 1990	border with West in 1990	no border with West in 1990
gender	0.53	0.50	0.52	0.51
age	32.29*	31.21*	31.11	31.82
spouse	0.76	0.73	0.75	0.73
university degree	0.05*	0.11*	0.09	0.09
vocational education	0.89	0.87	0.88	0.87
government sector	0.28*	0.37*	0.36	0.33
blue-collar employee	0.31*	0.36*	0.33	0.35
telephone	0.24	0.23	0.31*	0.20*
income, 1990	22758*	24973*	24576	23849

Notes: * difference in means significant at 5%. See footnote of Table 1.

Table 3: Second stage estimates: Migration

	Heckman's model		Nonparametric model	
	Migrants	Stayers	Migrants	Stayers
constant	6.02	6.61	7.95	6.63
	(1.286)	(0.231)	(1.399)	(0.266)
gender	0.74	0.38	0.72	0.37
	(0.125)	(0.022)	(0.125)	(0.024)
age	0.11	0.14	0.10	0.14
	(0.049)	(0.009)	(0.054)	(0.010)
age^2	-0.001	-0.001	-0.001	-0.001
	(0.0006)	(0.0001)	(0.0006)	(0.0001)
spouse	-0.35	-0.08	-0.35	-0.07
	(0.157)	(0.028)	(0.151)	(0.029)
university degree	0.57	0.49	0.53	0.47
	(0.221)	(0.046)	(0.240)	(0.045)
vocational education	-0.13	0.13	-0.21	0.13
	(0.197)	(0.038)	(0.192)	(0.048)
government sector	0.12	0.19	0.12	0.18
	(0.141)	(0.024)	(0.142)	(0.023)
blue-collar employee	-0.07	-0.10	-0.06	-0.09
	(0.158)	(0.023)	(0.124)	(0.024)
λ	0.67	0.25		
	(0.413)	(0.259)		
pscore			-5.86	-0.87
			(3.668)	(3.018)
pscore^2				52.05
				(56.98)
pscore^3				-378.62
				(305.96)
Observations	178	2865	177	2663
CM test 3rd moment		-0.00004		
		(0.0008)		
CM test 4th moment		0.0005		
		(0.0039)		

Note: standard errors, corrected for heteroskedasticity and for the first step generated regressors for Heckman's model and calculated as in Das et al (2003) for the nonparametrics model, are in parentheses. Dependent variable is log of the total annual average income. All time-variant demographic and human capital variables are of 1990. λ is the inverse Mills ratio. Pscore is estimated in the first stage propensity to move West. Covariates also include the state's unemployment rate and dummies for missing 1990 information. CM test refers to the conditional moment test for normality of Newey (1985), Pagan and Vella (1989). In the reported nonparametric model the intercept is estimated according to Andrews and Schafgans (1998).

Table 4: Second stage estimates: Commuting

	Heckman's model		Nonparametric model	
	Commuters	Stayers	Commuters	Stayers
constant	8.70	6.45	8.14	6.45
	(0.810)	(0.252)	(0.592)	(0.252)
gender	0.44	0.38	0.46	0.37
	(0.064)	(0.027)	(0.060)	(0.022)
age	0.06	0.15	0.06	0.15
	(0.025)	(0.010)	(0.025)	(0.010)
age^2	-0.0006	-0.001	-0.001	-0.002
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
spouse	-0.06	-0.08	-0.07	-0.07
	(0.068)	(0.029)	(0.072)	(0.028)
university degree	0.47	0.49	0.48	0.46
	(0.098)	(0.048)	(0.082)	(0.044)
vocational education	0.06	0.15	0.07	0.13
	(0.091)	(0.041)	(0.079)	(0.043)
government sector	-0.01	0.21	0.003	0.22
	(0.061)	(0.025)	(0.061)	(0.023)
blue-collar employee	0.01	-0.10	0.002	-0.09
	(0.064)	(0.025)	(0.058)	(0.023)
λ	-0.02	0.08		
	(0.134)	(0.130)		
pscore			3.88	
			(1.997)	
pscore^2			-9.54	
			(4.814)	
pscore^3				
Observations	430	2523	428	2431
CM test 3rd moment		-0.0040		
		(0.0020)		
CM test 4th moment		0.0115		
		(0.0057)		

Note: standard errors, corrected for heteroskedasticity and for the first step generated regressors for Heckman's model and calculated as in Das et al. (2003) for the nonparametrics model, are in parentheses. Dependent variable is log of the total annual average income. All time-variant demographic and human capital variables are of 1990. λ is the inverse Mills ratio. Pscore is estimated in the first stage propensity to move West. Covariates also include the state's unemployment rate and dummies for missing 1990 information. CM test refers to the conditional moment test for normality of Newey (1985), Pagan and Vella (1989). In the reported nonparametric model the intercept is estimated according to Andrews and Schafgans (1998).

Table 5: Intentions to treat effects and IV (LATE) estimates

	Intentions to treat: Move	Intentions to treat: Income	IV
	(1)	(2)	(3)
		A: Migration	
home owner, 1990	-0.039	0.011	
	(0.008)	(0.020)	
migrate			-0.273
			(0.538)
F-test on instrument in 1st stage		30.23	
		B: Commuting	
border with West, 1990	0.111	0.022	
	(0.015)	(0.022)	
commute			0.199
			(0.194)
F-test on instrument in 1st stage		62.52	

Note: robust standard errors are in parentheses. Panel A shows the estimates for migration, panel B - for commuting. Dependent variable in column 1 is migration or commuting dummy respectively, dependent variable in columns 2, 3, 4 is the log of average total annual income. Covariates include gender, age and its square, spouse indicator in 1990, educational and occupational dummies in 1990, state's unemployment rate in 1992 and dummies for missing 1990 information.

Table 6: Treatment effects for movers: Migration

Parametric	Nonparametric	LATE
A: Baseline model		
-0.19	0.40	-0.27
(0.531)	(0.233)	(0.538)
B: Robustness checks		
including unemployment in 1990		
-0.13	0.47	-0.25
(0.525)	(0.269)	(0.535)
including household income in 1990		
-0.04	0.32	-0.03
(0.521)	(2.461)	(0.524)
including telephone in 1990		
-0.16	-0.14	-0.24
(0.529)	(1.167)	(0.535)
including political interests in 1990		
-0.39	0.41	-0.38
(0.534)	(0.224)	(0.541)
including hours worked per week in 1990		
-0.16	0.37	0.09
(0.500)	(0.250)	(0.493)
excluding the self-employed		
-0.26	0.07	-0.32
(0.598)	(0.104)	(0.608)
retaining return and multiple migrants		
-0.22	0.07	-0.31
(0.482)	(0.090)	(0.497)
excluding "movers" from the control groups		
0.14	-0.52	0.03
(0.523)	(0.612)	(0.524)
including years for which the incomes are taken		
0.72	1.22	-0.54
(0.434)	(0.847)	(0.553)
C: Income growth as a dependent variable		
0.22	0.04	-0.11
(0.425)	(0.082)	(0.455)

Note: standard errors are in parentheses. Treatment effects are calculated as shown in Section 4. Dependent variable in the regressions is average annual total income in Panels A and B, and is income growth in Panel C . In the reported nonparametric model the intercept is estimated according to Andrews and Schafgans (1998).

Table 7: Treatment effects for movers: Commuting

Parametric	Nonparametric	LATE
A: Baseline model		
0.27	0.42	0.20
(0.230)	(0.029)	(0.194)
B: Robustness checks		
including unemployment in 1990		
0.31	0.42	0.23
(0.227)	(0.029)	(0.191)
including household income in 1990		
0.27	0.40	0.23
(0.228)	(0.029)	(0.192)
including telephone in 1990		
0.17	0.39	0.10
(0.236)	(0.072)	(0.202)
including political interests in 1990		
0.27	0.41	0.21
(0.228)	(0.029)	(0.193)
including hours worked per week in 1990		
0.31	0.40	0.32
(0.239)	(0.074)	(0.200)
excluding self-employed		
0.31	0.45	0.22
(0.235)	(0.030)	(0.197)
retaining return and multiple migrants		
0.30	0.40	0.25
(0.229)	(0.027)	(0.192)
excluding "movers" from the control groups		
0.24	0.43	0.19
(0.227)	(0.065)	(0.192)
including years for which the incomes are taken		
0.44	0.32	0.22
(0.184)	(0.027)	(0.187)
C: Income growth as a dependent variable		
0.13	0.29	0.26
(0.189)	(0.094)	(0.151)

Note: standard errors are in parentheses. Treatment effects are calculated as shown in Section 4. Dependent variable in the regressions is average annual total income in Panels A and B, and is income growth in Panel C . In the reported nonparametric model the intercept is estimated according to Andrews and Schafgans (1998).

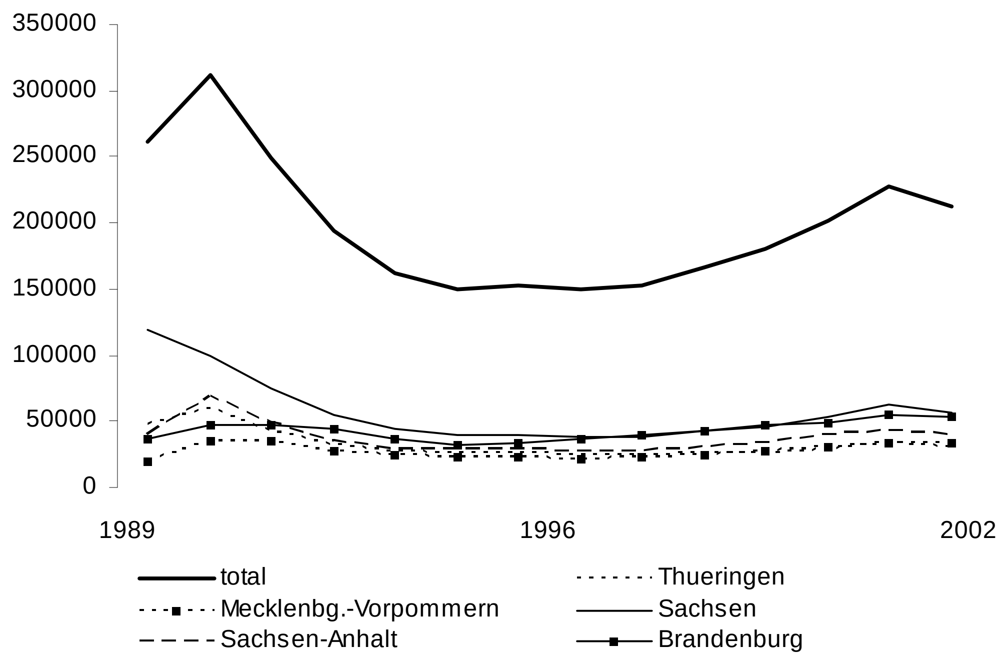

Figure 1: Emigration from East German länder to West Germany after the fall of the Berlin Wall. Source: numbers are from Heiland (2004). Note: East Berlin is omitted due to data unavailability.

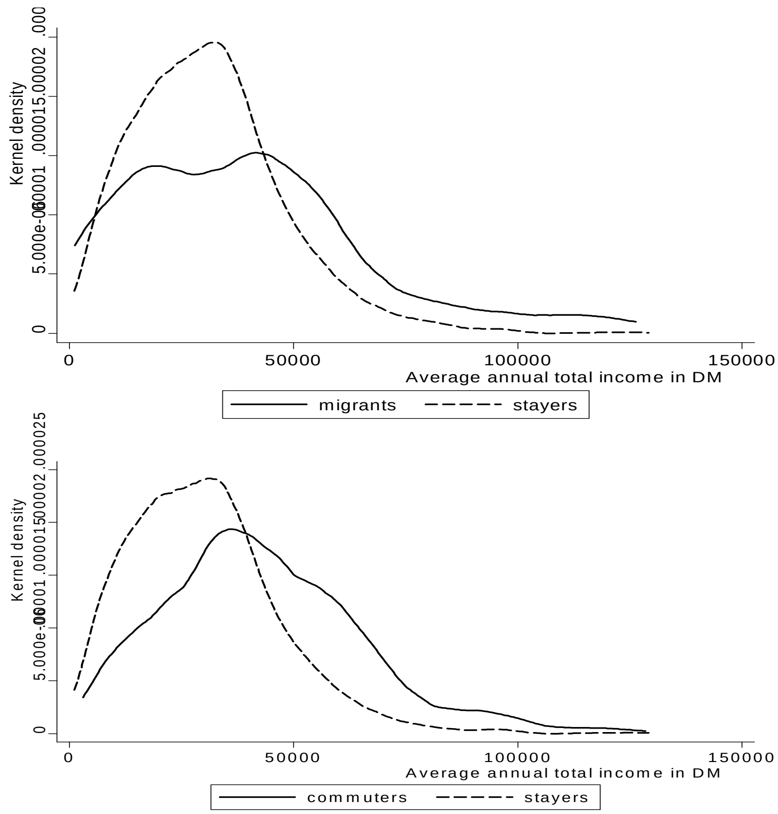

Figure 2: Kernel densities of the average annual total income for movers and stayers in Germany after unification. Source: GSOEP. Notes: see Section 2 for definitions.