Semiparametric Analysis of German East-West Migration Intentions: Facts and Theory *

Michael C. BURDA[1]

Wolfgang HÄRDLE[2]

Marlene MÜLLER[2]

Axel WERWATZ[2]

[1] Institut für Wirtschaftstheorie II, Wirtschaftswissenschaftliche Fakultät
Humboldt-Universität zu Berlin, Germany

[2] Institut für Statistik und Ökonometrie, Wirtschaftswissenschaftliche Fakultät
Humboldt-Universität zu Berlin, Germany

October 7, 1997

Abstract

East-West migration in Germany peaked at the beginning of the 90s although the average wage gap between Eastern and Western Germany continues to average about 25%. We analyze the propensity to migrate using microdata from the German Socioeconomic Panel. Fitting a parametric Generalized Linear Model (GLM) yields nonlinear residual behavior. This finding is not compatible with classical Marshallian theory of migration and motivates the semiparametric analysis. We estimate a Generalized Partial Linear Model (GPLM) where some components of the index of explanatory variables enter nonparametrically. We find the estimate of the nonparametric influence in concordance with a number of alternative migration theories, including the recently proposed option-value-of-waiting theory.

*The research in this paper was supported by Sonderforschungsbereich 373 at Humboldt-Universität zu Berlin, http://sfb.wiwi.hu-berlin.de. The paper is printed using funds made available by the Deutsche Forschungsgemeinschaft.

1 Introduction

German East-West migration has been the subject of several recent papers. Using microdata from the German Socio Economic Panel, Burda (1993), Büchel and Schwarze (1994) and Schwarze (1996) empirically investigate this issue. Especially interesting is the fact that, although migration peaked in the early 1990s following unification, the gap between average Eastern and Western wages remains about 25% as of 1997.

We take the empirical findings of Burda (1993) as our point of departure. We reanalyze the data by estimating a Generalized Linear Model (GLM) but find that the GLM does not provide a satisfactory fit. Estimating a semiparametric Generalized Partial Linear Model (GPLM) reveals a nonlinear influence of household income on the propensity to migrate form East to West. The functional form of this relationship can not be captured by a quadratic parametric fit.

We argue that the nonlinear influence of income on migration, while not implied by classical economic theory of migration, is compatible with the option value approach proposed by Dixit and Pindyck (1994) and applied recently to the migration decision by Burda (1995) and O'Connell (1997). In this approach migration is viewed as an investment with uncertain returns and irrecoverable fixed costs. Postponing migration means avoiding the fixed cost and observing part of the uncertain future while leaving the possibility of migrating open, thus implying an option value of postponing migration.

The remainder of the paper is organized as follows. In the following section we present a brief discussion of the classical (Marshallian) theory of migration behavior. In section 3 we introduce the data and discuss how facts and theory play together. Results from fitting a parametric GLM to the data are presented in section 4. As we shall see, standard Logit analysis does not appear to sufficiently capture the phenomenon underlying the observations. We therefore turn to a more flexible setting by allowing some components to take a nonparametric form. These semiparametric Generalized Partial Linear Models (GPLM) are described and estimated in sections 5. In section 6 we discuss our findings in the light of option value theory. Section 7 concludes the paper.

2 Some Theoretical Considerations

Since Ravenstein's pathbreaking work on the determinants of migration more than a century ago, income has been the focus of economists' attempts to explain spatial mobility. More precisely, the difference between income at home (W^E) and the income attainable by migrating (W^W) has been singled out as the key explanatory variable.

A forward-looking agent will not only care about the current income differential (which we assume is known) but also about future income differentials. That is, he will consider the expected present value (net of the income stream obtainable from not-migrating) of the income stream he will receive if he decides to migrate.

But even if this expected present value is positive the agent may not migrate if the fixed costs of migrating are sufficiently high. Such fixed costs will include pecuniary components associated with physically moving a household from one place to another. In addition, moving away means leaving behind an environment one was accustomed to as well as friends and family members.

Following classical ("Marshallian") economic theory, we may therefore say that a rational, forward-looking agent will migrate if the expected present value of the income stream from migrating exceeds the fixed cost, or if the expected net present value from migrating (net of fixed costs) is positive. Incorporating risk aversion will change the trigger rule, but at most by a constant amount which would depend on the relative riskiness of the options and individual preferences.

Under a number of weak assumptions about the stochastic process generating relative income, the expected present value of future gains from migration will be a function of the current observed income differential, and for plausible assumptions this relationship will be linear. For instance, if absolute per-period, East-West income differential $\Omega_t = W_t^W - W_t^E$ follows an arithmetic Brownian process with negatived drift ν, then the expected present value of migration is given by $V^m = (\Omega_0 - \nu/\delta)/\delta$, where δ denotes the discount rate and Ω_0 the current income differential. [1]

Let the fixed monetary costs (including monetary equivalent of utility loss from moving) be given by F or f and denote the migration decision by the binary variable Y ($Y = 1 \longrightarrow$ migration). Then the decision rule for a rational agent can be formally written as:

$$\begin{aligned} Y &= 1 \quad \text{if } V^m = \frac{1}{\delta}(\Omega_0 - \nu/\delta) - F > 0 \\ Y &= 0 \quad \text{otherwise} \end{aligned} \tag{1}$$

Figure 1 shows the prediction of the theory. The dashed straight line depicts the net present value of migrating, V^m, as a function of the current income differential. The slope of the straight line is given by $1/\delta$. If the income differential equals zero the

[1] Alternatively, suppose that the relative differential $\omega_t = (W_t^W - W_t^E)/W_t^W \approx \ln(W_t^W) - \ln(W_t^E)$ follows a geometric Brownian motion with negative drift μ. Then we can write $d\omega_t = -\mu\omega t + \sigma\omega t dz$ where dz is a Wiener diffusion process. In this case the expected present discounted value of relative differentials at $t = 0$ is simply $\omega_0/(\mu + \delta)$. In case of a known finite rest of life T, the value is $\omega_0(1 - \exp(\mu + \delta))/(\mu + \delta)$.

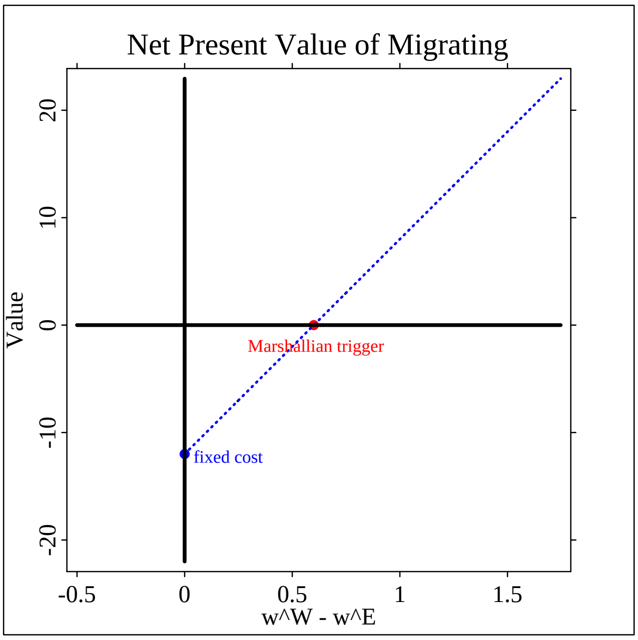

Figure 1. Marshallian decision rule for migrating as a function of the wage gap

net present value from migrating is essentially equal to the negative value of the fixed costs (point labeled "fixed cost"). If the income differential equals the "Marshallian trigger" then the net present value is exactly zero. Any income differential exceeding this trigger implies migration while income differentials smaller than the trigger imply the opposite.

The theory delivers a clear prediction that an increase in current income will decrease migration propensity for a given set of alternatives available in the West. This is depicted is Figure 2 which graphs the net present value of migrating as a function of the current income in the East (W_0^E).

3 The Data

In the empirical analysis we use data drawn from the German Socio Economic Panel (GSOEP). The GSOEP is a representative panel survey of German households that was extended to the former East in 1990. We use 3367 observations from the GSOEP's second East-German wave which was collected in the spring of 1991 (time $t = 0$). All calculations were carried out with the statistical computing environment XploRe.

In the wave of the GSOEP considered, there are only a few actual migrants. We therefore use migration propensity ("intention") as the dependent variable Y. The theoretical discussion of the previous section has focused on the income differential between host region and home region and the fixed cost of migrating as the key ex-

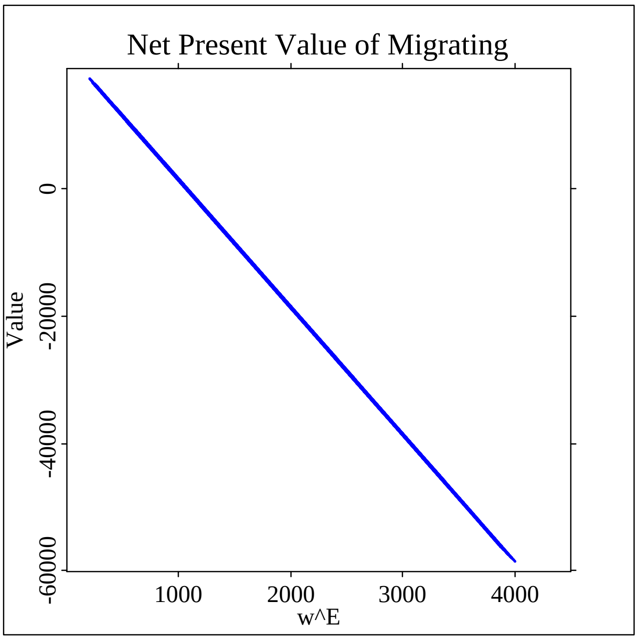

Figure 2. NPV of migrating as a function of W_0^E

planatory variables. Yet, measuring both quantities poses a challenge. Regarding the income differential, we are faced with the problem that the potential income in the West is not observable. Hence, some imputation is generally necessary.

Since Germany shares the same institutions and language one could assume that upon migration eastern Germans are able to employ at least some component of their human capital, earning "western returns" for their attributes, at least up to a (macroeconomic) constant.

One natural approach to estimating W_0^W or w_0^W is to employ estimates of a traditional earnings equation of the Mincer type, which attributes observed wages to either market "returns" multiplied by observable measures of human capital endowment (education, experience, training, tenure) or to attributes unobservable to the econometrician modeled as a random disturbance.

But estimating this relation on a sample of Westerners will most likely produce estimates that will seriously suffer from selection bias. Moreover, it is unclear how to use these estimates to calculate an imputed Western wage for those Easterners who are registered as unemployed or out of the labor force. Rather than producing spurious findings based on biased estimates of the West-East income differential we decided to include income in the East only. We shall discuss the observed facts, though, as a function of the income differential in section 6.

The GSOEP data provides a multitude of variables that arguably are related to the intention to migrate from the East to the West. Starting from a set of roughly 30 potential explanatory variables considered in the empirical analysis of Burda (1993) we used economic intuition and statistical selection criteria to limit the number of explanatory variables. This was merely done for better exposition of the facts. The proposed statistical method is valid for any dimension of the vector of explanatory variables.

		Mean	S.D.	expected effect
Y	**migration intention**	.394	.489	
X_1	**female**	.511	.500	
X_2	**partner**	.854	.353	–
X_3	**owner**	.322	.467	–
X_4	**family/friends in west**	.855	.352	+
X_5	**unemployed/jobloss certain**	.196	.397	+
X_6	**environmental satisfaction**	3.9	2.4	–
X_7	**city size < 10,000**	.522	.499	
X_8	**city size 10–100,000**	.342	.474	
X_9	**university degree**	.085	.278	
X_{10}	**age**	39.4	12.8	–
	min: 18, max 65			
X_{11}	**household income**	2189.5	754.7	
	min: 200, max: 4000			

Table 1. Summary Statistics

Summary statistics for Y and the explanatory variables are given in Table 1. Presence of a partner, home-ownership and increasing age are expected to increase the fixed cost of migrating whereas relatives or friends in the West supposedly have the opposite effect. Age will also influence the migration decision via the discount rate. The variable *environmental satisfaction* is measured on a scale from 1 ("very unhappy with environmental conditions") to 10 ("very happy ") and can therefore be expected to have a negative influence on migration propensity. The sign of the coefficients of the gender, city size and education variables is rather unclear apriori.

We have separated **age** and **household income** from the remaining explanatory variables in the table as -for the purposes of this study- they can be regarded as *continuous* explanatory variables.

4 Parametric Estimation Results

Collect the explanatory variables described in the previous section into the vector x. The goal of the empirical analysis is to estimate the probability of migration intention, i.e. $E(Y|x) = Prob(Y = 1|x)$. A natural starting point for estimating this probability is fitting a parametric GLM. More precisely, we estimated a Logit model.

Although this model is well known we briefly discuss it here. This is helpful in contrasting it with the semiparametric model to be discussed in section 5.

The parametric Logit model is based on two assumptions :

- **Latent-variable assumption**

$$\begin{aligned} Y &= 1 \quad \text{if} \quad Y^* = x^T\beta - u > 0 \\ Y &= 0 \quad \text{otherwise} \end{aligned} \tag{2}$$

  That is, underlying the observable binary dependent variable Y is an unobserved, latent variable Y^*, assumed to be the sum of a linear index of the explanatory variables x (common to all individuals in this study) and an individual error term u Here β is a vector of unknown coefficients that has to be estimated from the data.

- **Distributional assumption**
  Let $F_{u|x}(\bullet)$ denote the cumulative distribution function (cdf) of u conditional on x. The Logit model assumes that $F_{u|x}$ is the logistic distribution function for all x.

Combining both assumptions gives

$$E(Y|x) = Prob(Y = 1|x) = \{1 + \exp(-x^T\beta)\}^{-1} \tag{3}$$

As usual, $G(u) = \{1 + \exp(-u)\}^{-1}$ is called the (inverse) link function.

Table 2 gives the Maximum Likelihood Logit estimates of β. Most coefficients have the expected sign: age, a partner, home ownership and environmental satisfaction reduce migration propensity whereas family or friends in the West and poor labor market prospects in the East have the opposite effect.

The estimated coefficient of the linear logit specification suggests that migration propensity significantly increase with household income. Figures 3 and 4 reflect the actual dependence of the response Y on variables age and income. We have plotted each variable versus the logits

$$\text{logit} = \log\left(\frac{\widehat{p}}{1 - \widehat{p}}\right)$$

where $\widehat{p}$ are the relative frequencies for $Y = 1$ (migration intention). Essentially, these logits are obtained from classes of neighbored realizations (where the range of either age or income has been divided into 50 equidistant intervals). In case that $\widehat{p}$ was 0 or

dependent variable: migration intention		
Variable	estim. coeff.	t-ratio
cons	1.864	7.74
female	-.233	-3.03
partner	-.325	-2.87
owner	-.576	-5.79
family/friends in west	.647	5.61
unemployed	.217	2.24
environmental satisfaction	-.057	-3.52
city size < 10,000	-.718	-5.69
city size 10-100,000	-.347	-2.91
university degree	.481	3.56
age	-.050	-14.89
household income	.0001202	2.22
sample size: 3367, log likelihood: -1992.7		

Table 2. Logit Estimates

1, several classes have been joint. Thicker bullets correspond to more observations in a class.

Figure 3 shows that age has an almost linear influence on migration intention, whereas the relationship between income and migration intention follows a slightly U–shaped curve. Economic theory stresses the importance of income as an explanatory variable. From the perspective of building a satisfactory statistical model income – being a continuous variable – should be entered in a nonlinear way.

If we include the square of household income as an additional regressor then both income coefficients are individually insignificant. This finding may lead an analyst to conclude that income does not have a nonlinear influence. Yet, if we add income cubed as a regressor to the model that already includes income and income squared then all three income coefficients are individually as well as jointly significant. These findings are summarized in Table 3. It should be noted, though, that the significance level for the cubic model was not Bonferroni corrected to incorporate the fact that we previously reject the quadratic model.

Rather than continuing with the refinement of this parametric specification we decided to estimate a semiparametric Generalized Partial Linear Model which allows the data to freely determine the shape of the influence of income on migration propensity. By means of generalized additive modeling this can be extended to the variable age as well. An analysis of this model yielded a linear dependence of migration propensity on age (as in Figure 3). We therefore included only income as a possible nonlinear candidate.

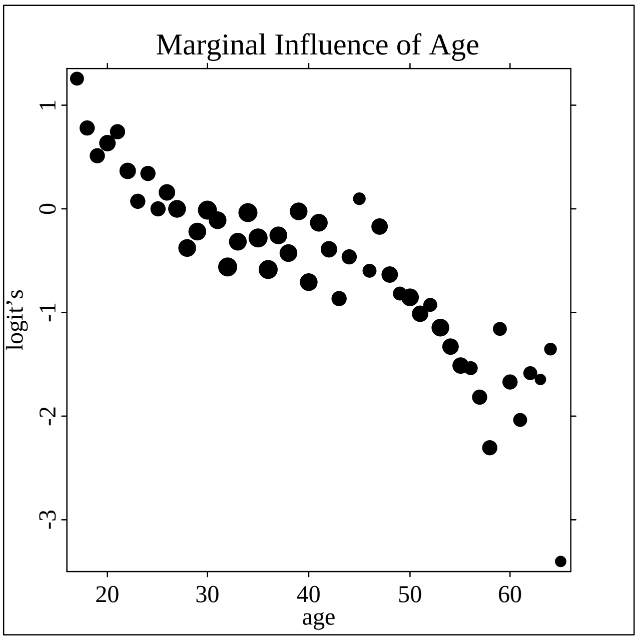

Figure 3. Marginal influence of age on migration intention, visualized by logits on classes of age

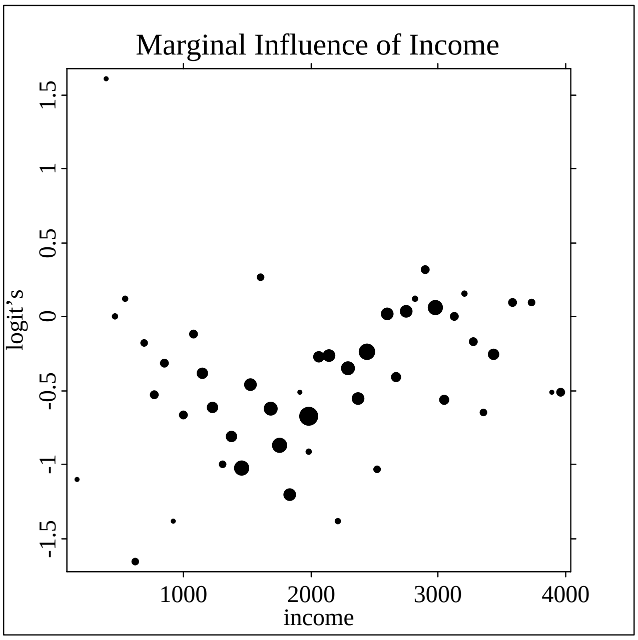

Figure 4. Marginal influence of income on migration intention, visualized by logits on classes of income

5 Semiparametric Estimation Results

Before turning to estimates, we will briefly introduce the generalized partially linear model (GPLM). The GPLM assumes that the mean of Y is related to an index of

Variable	estim. coeff.	t-ratio
"quadratic" model		
household income	-.0001288	-0.507
household income^2	5.46e-08	1.002
"cubic" model		
household income	-.0016491	-2.130
household income^2	8.08e-07	2.206
household income^3	-1.12e-10	-2.080
dependent variable: migration intention same regressors as above besides income		

Table 3. Parametric specification search

explanatory variables via a known link function G. In the particular GPLM used below we will take G to be the distribution function of the logistic distribution, i.e. $G(\bullet) = 1/\{1 + \exp(-\bullet)\}$.

Contrary to the Logit model of the previous section the index of explanatory variables is comprised of a linear parametric component and a nonparametric component. That is, the GPLM assumes that

$$E(Y|x,t) = G\{x^T\beta + m(t)\}. \tag{4}$$

where –in a slight abuse of notation– we have collected the explanatory variables that enter the argument of $G(\bullet)$ linearly in the $p \times 1$ vector x, and those that enter nonlinearly in the $q \times 1$ vector of variables t. The unknown quantities that need to be estimated are the parameter vector β and the unknown function $m(\bullet)$. Note that there is no intercept parameter since it can be absorbed into the nonparametric part $m(t)$. In the empirical analysis x will – with the exception of age – be made up of discrete (categorical) variables while t solely contains household income.

The estimation methods for model (4) are based on the idea that an estimate $\widehat{\beta}$ can be found for known $m(\bullet)$, and an estimate $\widehat{m}(\bullet)$ can be found for known β. In what follows we will concentrate on *profile likelihood* estimation which goes back to Severini and Wong (1992), Severini and Staniswalis (1994).

Denote by $L(\mu, y)$ the individual log–likelihood, where $\mu = E(Y|x,t) = G\{x^T\beta + m(t)\}$. The profile likelihood uses two different likelihood functions for the estimation of the parametric and semiparametric components. The usual likelihood for n i.i.d. observations (x_i, t_i, y_i)

$$\mathcal{L}(\beta) = \sum_{i=1}^{n} L\left\{\beta^T x_i + m_\beta(t_i); y_i\right\} \tag{5}$$

is used to obtain $\widehat{\beta}$ and a "smoothed" likelihood

$$\mathcal{L}_h(\eta) = \sum_{i=1}^{n} K_h(t - t_i)\, L\left(\beta^T x_i + \eta; y_i\right) \tag{6}$$

for the nonparametric smooth function $\widehat{m}_\beta(t) = \eta$ at point t.

The computational algorithm consists in searching maxima of both likelihoods simultaneously. A detailed description of the algorithm can be found in the Appendix. It turns out that the resulting estimator $\widehat{\beta}$ is $\sqrt{n}$–consistent and asymptotically normal, and that estimators $\widehat{m} = \widehat{m}_{\widehat{\beta}}$ are consistent in supremum norm, see Severini and Staniswalis (1994).

Table 4 gives the GPLM estimates of β in a model that includes the same explanatory variables as the Logit fit of Table 2. The Logit estimates and their t-ratios are also reported to conveniently compare results across the different approaches. In general,

dependent variable: migration intention				
	GPLM estimates		Logit estimates	
Variable	coeff.	t-ratio	coeff.	t-ratio
female	-0.238	-3.09	-.233	-3.03
partner	-0.282	-2.44	-.325	-2.87
owner	-0.569	-5.71	-.576	-5.79
family/friends in west	0.640	5.54	.647	5.61
unemployed	0.216	2.23	.217	2.24
environmental satisfaction	0.056	-3.47	-.057	-3.52
city size < 10,000	-0.689	-5.43	-.718	-5.69
city size 10–100,000	-0.323	-2.71	-.347	-2.91
university degree	0.471	3.48	.481	3.56
age	-.050	-14.89	-.050	-14.89
sample size: 3367, log likelihood: -1989.8, GPLM bandwidth: 0.3				

Table 4. GPLM Estimates

the GPLM estimates are very close to their Logit counterparts.

In terms of the GPLM, income plays the role of the variable t in (4). The estimated influence of income is depicted in Figure 5, with income on the horizontal axis and the estimate of $m(t)$ on the vertical axis. The highly nonlinear estimate of $m(t)$ strongly contrasts with the linear influence of income implied by the Logit model which we have also included in Figure 5.

The GPLM fit suggests a U-shaped influence over the range of income values that carry most of the mass of the income distribution. The bandwidth h underlying the estimate of $m(t)$ was set equal to $h = 0.3$ but a U-shaped estimate is obtained for a range of values of h.

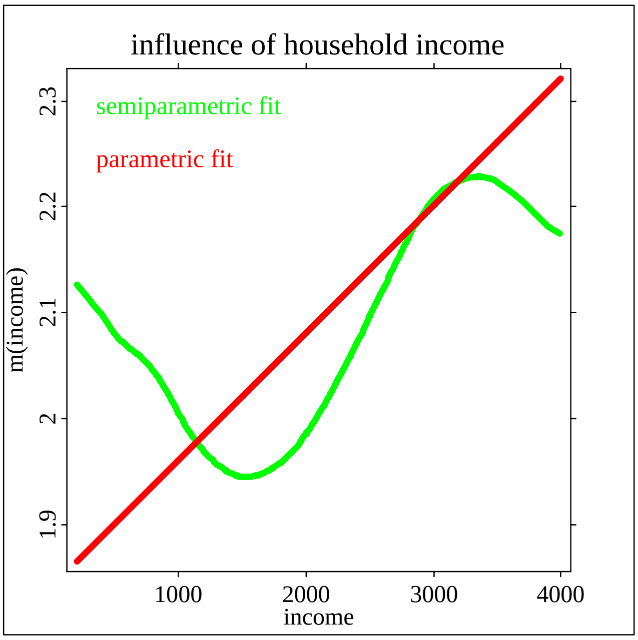

Figure 5. Influence of the net household income on migration propensity

The visual impression of Figure 5 suggests that the estimate of $m(t)$ significantly deviates from the estimated linear influence of the parametric GLM fit. We use a test procedure to formally test that $m(t)$ is a linear function:

$$H_0 \; : \; m(t) = \alpha\, t + \alpha_o\,,$$
$$H_1 \; : \; m(t) \text{ is an arbitrary smooth function,}$$

This test is based on comparing the semiparametric estimates with the parametric estimates

$$(\widetilde{\beta}, \widetilde{\alpha}, \widetilde{\alpha}_o) = \arg \min_{\beta,\alpha,\alpha_o} \sum_{i=1}^{n} L\left[G\{x_i^T\beta + \alpha\, t_i + \alpha_o\}\,;y_i\right], \tag{7}$$

where α denotes the coefficient of income and α_o the constant in the parametric fit.

A test of the hypothesis GLM (logit model) against the alternative of a GPLM may be based on the likelihood ratio statistic. Denote by $\widetilde{\mu}_i = G(x_i^T\widetilde{\beta} + \widetilde{\alpha}\, t + \widetilde{\alpha}_o)$ the parametric GLM fit and by $\widehat{\mu}_i = G\{x_i^T\widehat{\beta} + \widehat{m}(t)\}$ the GPLM fit. Hastie and Tibshirani (1990) propose to use

$$R = 2\sum_{i=1}^{n} L(\widehat{\mu}_i, y_i) - L(\widetilde{\mu}_i, y_i) \tag{8}$$

which has heuristically a distribution that is similar to a χ^2 distribution. However, the degrees of freedom for the GPLM need to be replaced by an approximate value and theoretic distribution of R is unknown.

Härdle, Mammen and Müller (1996) propose a modification of the test statistic R. This modification is based on the fact that a direct comparison of $\widehat{m}(t)$ and $\widetilde{\alpha}\, t + \widetilde{\alpha}_o$

can be misleading because $\widehat{m}$ has a non–negligible smoothing bias. This holds even under the linearity hypothesis. Hence, a bias–corrected parametric estimate $\overline{m}(t)$ is used instead of $\widetilde{\alpha}\, t + \widetilde{\alpha}_o$.

Using this bias–corrected $\overline{m}(t)$ the following modified likelihood–ratio test statistic is computed

$$R^M = 2 \sum_{i=1}^{n} L(\widehat{\mu}_i, \widehat{\mu}_i) - L(\overline{\mu}_i, \widehat{\mu}_i), \tag{9}$$

where $\overline{\mu}_i = G\{x_i^T \widetilde{\beta} + \widetilde{m}(t_i)\}$ is the bias corrected GLM fit and $\widehat{\mu}_i$ the GPLM fit as before.

Härdle et al. (1996) show asymptotic normality of R^M. The proof of this result is based on showing that the asymptotic expansion of R^M behaves approximately like a sum of $O(h)$ independent summands. This is typically not very large and indeed simulations show that the normal approximation need not work well for R^M (Müller, 1997). Therefore, for the calculation of quantiles, it is recommended to use the the following bootstrap procedure:

1. Generate samples $\{Y_1^*, \ldots, Y_n^*\}$ under the parametric hypothesis with $E^*(Y_i^*) = G(x_i^T \widetilde{\beta} + \widetilde{\alpha}\, t_i)$. Here E^* and denotes the conditional expectation given $(x_1, t_1, \ldots, x_n, t_n)$.
2. Calculate estimates $\widehat{\beta}^*, \widehat{m}^*, \widetilde{\beta}^*, \widetilde{\alpha}^*, \widetilde{m}^*$ based on the bootstrap samples $\{(x_1, t_1, Y_1^*)$, …, $(x_n, t_n, Y_n^*)\}$. Furthermore, calculate test the statistic R^{M*}. Repeat this n^* times. The quantiles of the distribution of R^M can be estimated by the quantiles of the conditional distribution of R^{M*}.

Since in our case the distribution of Y is completely specified by $EY = \mu = G(x^T\beta + \alpha\, t + \alpha_o)$ (under linearity hypothesis) we resample from the Bernoulli distribution with parameters $\widetilde{\mu}_i = G(x_i^T \widetilde{\beta} + \widetilde{\alpha}\, t_i + \widetilde{\alpha}_o)$ (the parametric GLM fit).

h	0.1	0.2	0.25	0.3	0.4
R	0.028	0.021	0.019	0.017	0.016
R^M	0.053	0.069	0.130	0.269	0.602
R^{M*}	0.015	0.005	0.005	0.005	0.010

Table 5. Observed significance levels for linearity test for migration data, $n = 3367$. 200 bootstrap replications. Bandwidth h in % of range of household income

Table 5 shows the result of both test procedures for the GLM vs. the GPLM. With R^M we denote the test using test statistic (9), where the test has been carried out using the normal approximation. R^{M*} bootstrap denotes the results for the bootstrapped quantiles of R^M. Since an optimal bandwidth choice for the GPLM is not known, all tests were performed for a sequence of bandwidths. However, we can recognize a clear rejection of the linearity hypothesis across all bandwidths for the R and the bootstrapped R^{M*}. The normal approximation for R^M works bad for higher bandwidth levels as was already indicated above.

6 Explaining the Results: Alternative Theories

In the previous section we have found a significant nonlinear relationship between migration propensity. This is at variance with the linear relationship implied by the classical theory of migration outlined in section 2. In this section we will briefly outline theoretic models of migration that may rationalize the shape of the estimate of Figure 5.

Option Value Theory

One limiting aspect of the Marshallian theory of migration of section 2 is its "all-or-nothing" aspect; either migration occurs now or never. The work of Dixit and Pindyck (1994) and others has shown that postponement of the decision without forsaking it can be a valuable option under a large class of irrevocable investment problems. Heuristically, if the agent has the ability to delay a decision, he or she can acquire more information and increase the likelihood that the decision will not be regretted in the future. Following Burda (1995) we will outline how these option-value arguments may be applied to the migration problem.

In section 2 we derived classical economics' rule for the migration decision:

$$\begin{aligned} Y &= 1 \qquad \text{if } V^m = \frac{1}{\delta}\left(\Omega_0 + \nu/\delta\right) - F > 0 \\ Y &= 0 \qquad \text{otherwise} \end{aligned} \tag{10}$$

Migrating today means incurring the fixed cost F and forgoing the current and future income in the sending region. This opportunity cost of migrating is incorporated in (10) since $\gamma\,\Omega_0$ is the expected present value of the income stream from migrating net of the income stream obtainable by not-migrating.

Migrating today, however, also means forgoing the opportunity to postpone migration. This opportunity has positive (expected) value today because waiting brings more information about the future while it leaves open the possibility to still migrate should the future evolve favorably (or not to migrate in the unfavorable case).

We will denote this opportunity cost of migrating today as V^o and refer to it as the *option value of waiting*. Certainly, it will be a function of the current wage differential $\Omega_0 = W_0^W - W_0^E$. V^o is equal to what one is willing to pay for the option to postpone the migration decision rather than having to decide now or never. It can be calculated as the difference between the expected net present value from postponing migration, V^p, and the expected net present value from migrating today, V^m. See Dixit and Pindyck (1994) for several instructive numerical examples. We give a graphical illustration in Figure 6.

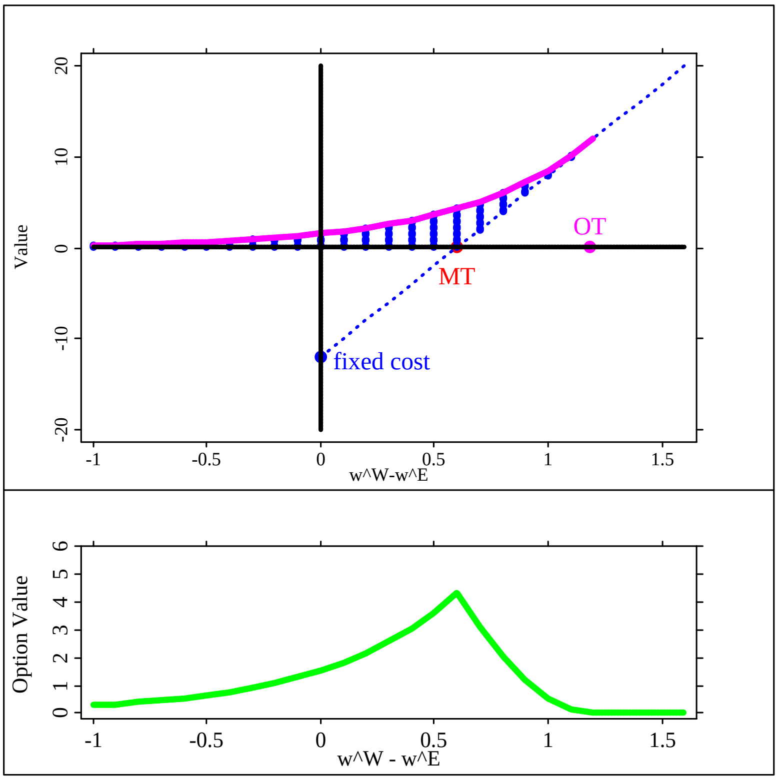

Figure 6. The Option Value of Waiting

Figure 6 graphs V^o (kinked curve in the lower panel), V^p (the positively sloped curve in the upper panel) and V^m (the dashed straight line in the upper panel) as functions of the current income differential.

If the current wage differential is below MT (the "Marshallian trigger") immediate migrating does not have positive net value ($V^m < 0$). Hence, V^o, –which is the amount a rational agent is willing to pay for the option to postpone investment– is just equal to V^p.

If the current wage differential is between MT and OT ("option-value trigger") then immediate migration has positive expected value and hence $V^o = V^p - V^m$. We have displayed the values V^o as vertical bars in the upper panel for selected values of the current wage differential.

If the current wage differential is above OT then V^o is zero: the current wage differential is so large that any further postponement of migration has zero value.

It appears from Figure 6 that V^o has the opposite shape as the estimated relationship of the previous section. But V^o is the option value of **postponing** migration. That is, high values of V^o imply a low propensity to migrate and vice versa. This is clearly evident if we rewrite the "classical" decision rule (10) to incorporate the option value of waiting:

$$\begin{aligned} Y &= 1 \quad \text{if } \frac{1}{\delta}(\Omega_0 + \nu/\delta) - F - V^o(\Omega_0) > 0 \\ Y &= 0 \quad \text{otherwise} \end{aligned} \tag{11}$$

As a consequence, to graphically see the implication of the option value theory on executing the migration option we have to flip around V^o which produces a U-shaped relationship. This has been done in Figure 7.

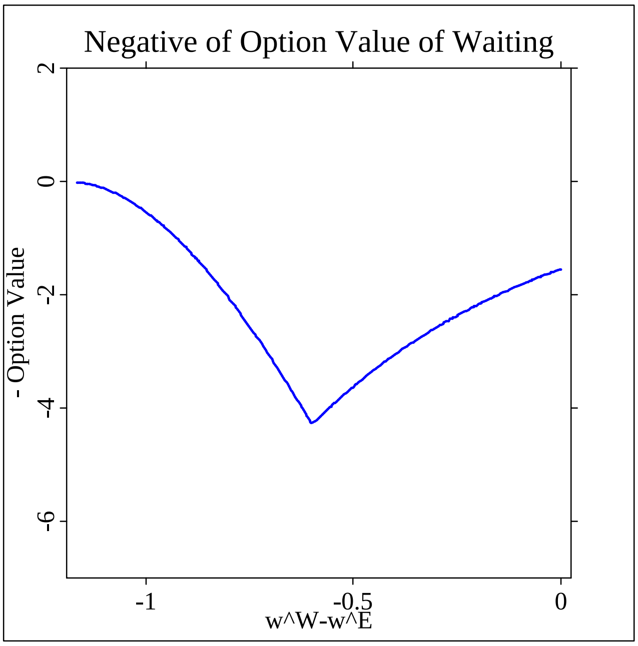

Figure 7. Option Value of Waiting as an opportunity cost of migrating now

One may raise the objection to the previous discussion that it is arguing in terms of the income *differential* while the empirical analysis is employing income in the East only. Figures 8 and 9 try to clarify this point.

The top panel of Figure 8 is a repetition of the lower panel of Figure 6. It plots the option value of waiting against the West-East income differential. The middle panel of Figure 8 plots the (hypothetical) Western income (vertical axis) versus the Eastern income. The lower straight line is the 45 degree line whereas the upper straight line corresponds to the hypothesis that the Western income is proportionally higher for each given level of Eastern income. Now suppose that the option value of postponing migration is depending on the income differential as depicted in the top panel of Figure 8. Then, under the hypothesis of the middle panel, the option value of postponing

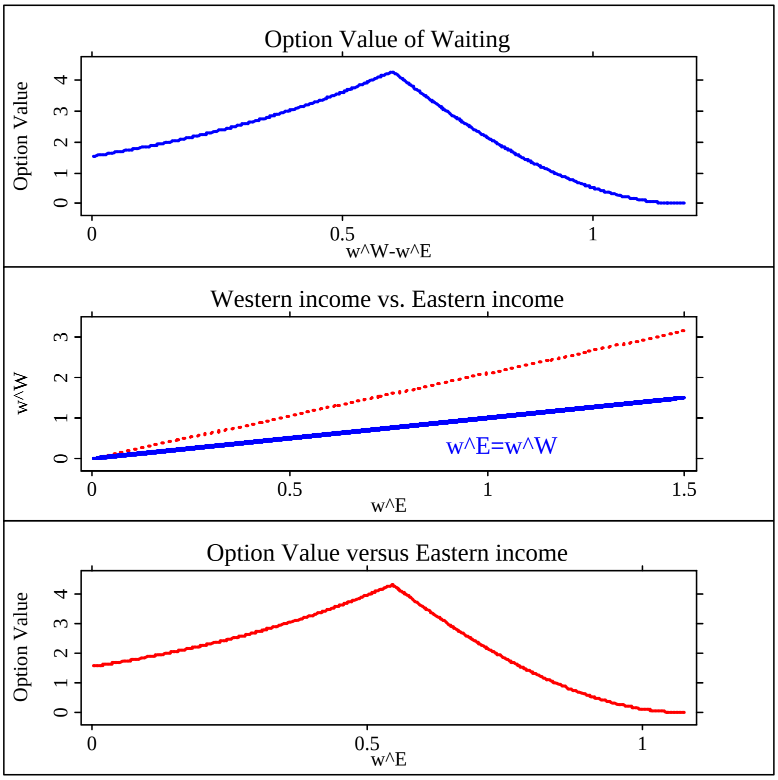

Figure 8. Eastern income versus Western income

migration plotted as a function of the income in the East (lower panel) has the same shape as if it is plotted as a function of the income differential.

Similarly, Figure 9 shows that different hypothesis about the relationship between Eastern and Western income still preserve the nonlinearity of the option value – regardless whether it is plotted as a function of the income differential or income in the East. Specifically, the parabola in the middle panel of this figure reflects the hypothesis that Easterners with a low income (expect to) receive a relatively high Western income, those with a mid-range income receive a rather small increase in the West and individuals with a high Eastern income expect a relatively strong increase in income by moving to the West. Under this assumption about the relationship between income in the East and income in the West, and under the assumption that the option value of waiting depends on the current West–East income differential as depicted in the top panel of Figure 9, we obtain the nonlinear relationship between the option value and income in the East as shown in the lower panel of Figure 9.

While the previous discussion has demonstrated the ability of option value theory to rationalize the estimated relationship between income and migration propensity, it has by no means incorporated all theoretical aspects of the migration decisions. In the remainder of this section we will therefore briefly discussion some of the issues that have been ignored up to this point.

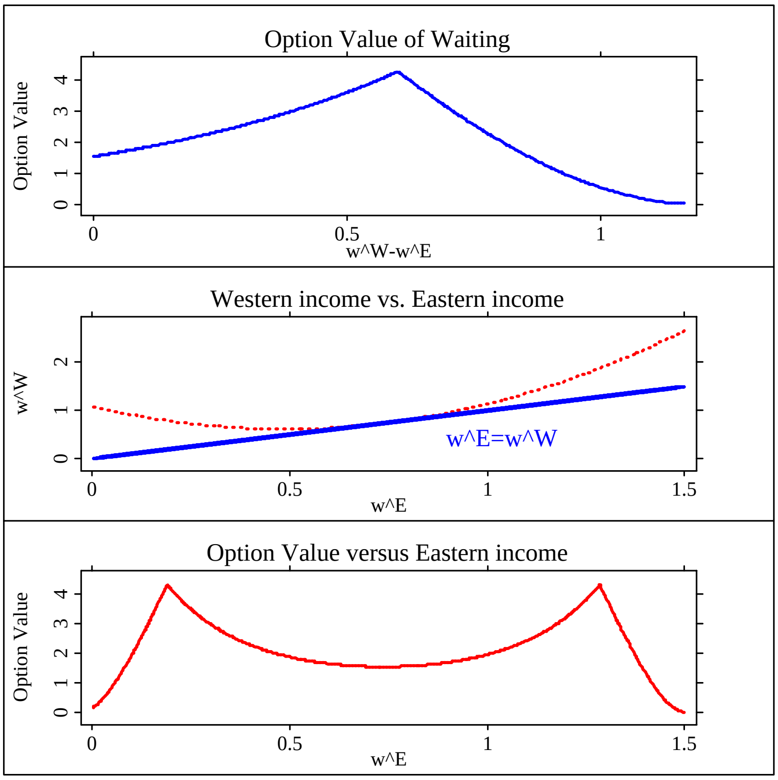

Figure 9. relationship between Eastern and Western income

Risk–aversion, Income effects and the demand for immobility

In the previous discussion, there is no mention of risk aversion nor the possibility that increasing wealth or income could increase the demand for immobility or mobility, depending on the utility function. This hypothesis has been put forward and investigated by, among others Faini and Venturini (1993) and Faini and Venturini (1994). Assuming current place of residence is a normal good, the income effect of higher absolute wages at home implies a lower propensity to migrate. (Alternatively, wealthier individuals may seek to flee their in-laws by moving, reducing dependence on relatives, etc.)

In general, curvature in the utility function (as opposed to strict linearity in previous sections) will lead to a reduced valuation of the migration decision if the primary source of uncertainty is in income abroad. An exception is Stark (1989) who shows that in some cases migration may serve a function of risk diversification or reduction. Below we show an example of how introducing curvature in the utility function (risk aversion, decreasing marginal utility) could affect the valuation of the migration decision without considering any option value. In the net, this reasoning predicts either a negative or a positive effect of absolute income on migration propensities.

Borrowing constraints and liquidity effects

Suppose that a component of moving costs F, realistically, must be paid in cash, and cannot be financed out of future earnings in the host country. In such a situation, the absolute value of current income (and not relative to abroad) matters for some range – when assets are inadequate to finance the move. When the wage rises, some households which may have been willing to migrate for some time can do so, financing the move out of current income. This reasoning predicts a positive effect of home wage/income on migration propensity for some range of current income.

7 Conclusions

In this paper we have empirically analyzed the propensity to migrate using microdata from the German Socioeconomic Panel. Fitting a parametric Generalized Linear Model (GLM) did not produce a satisfactory estimate of the influence of income. By estimating a Generalized Partial Linear Model (GPLM) we found a U-shaped relation between income and (the systematic part of) migration propensity. This functional form was not detected by a specification search within the framework of a parametric GLM.

We have argued that the estimated influence may be explained by a number of alternative determinants of migration, including the recently proposed option-value-of-waiting theory, liquidity constraints, wealth-conditioned immobility, as well as unobservable heterogeneity.

8 Acknowledgments

We would like to thank Swetlana Schmelzer for her help in designing the graphs in XploRe. The participants of the Berlin–Paris seminar made helpful suggestions. Alan Kirman pointed us to the second theoretical explanation.

References

Büchel, F. and Schwarze, J. (1994). Die Migration von Ost- nach Westdeutschland – Absicht und Realisierung, *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* **27**(1): 43–52.

Burda, M. (1993). The determinants of east-west german migration – some first results, *European Economic Review* **37**: 452–461.

Burda, M. C. (1995). Migration and the option value of waiting, *Economic and Social Review* **27**: 1–19.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment Under Uncertainty*, Princeton University Press, Princeton.

Faini, R. and Venturini, A. (1993). Trade, aid and migrations, *European Economic Review* **37**: 435–442.

Faini, R. and Venturini, A. (1994). Migration and growth: The experience of southern europe, *Discussion Paper 964*, CEPR.

Härdle, W., Mammen, E. and Müller, M. (1996). Testing parametric versus semiparametric modelling in generalized linear models, *SFB 373 Discussion Paper 960028*, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*, Vol. 43 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, London.

Müller, M. (1997). Computer–assisted generalized partial linear models, *Interface'97 Proceedings*, Houston, Texas.

O'Connell, P. (1997). Migration under uncertainty: 'try your luck' or 'wait and see', *Journal of Regional Science* **47**: 331–347.

Schwarze, J. (1996). Beeinflußt das Lohngefälle zwischen Ost– und Westdeutschland das Migrationsverhalten der Ostdeutschen ? *Allgemeines Statistisches Archiv* **80**(1): 50–68.

Severini, T. A. and Staniswalis, J. G. (1994). Quasi-likelihood estimation in semiparametric models, *Journal of the American Statistical Association* **89**: 501–511.

Severini, T. A. and Wong, W. H. (1992). Generalized profile likelihood and conditionally parametric models, *Annals of Statistics* **20**: 1768–1802.

Stark, O. (1989). *The Migration of Labor*, Basil Blackwell, Oxford.

XploRe (1997). *XploRe – the interactive statistical computing environment*, WWW: http://www.xplore-stat.de.

Appendix: Algorithm for GPLM

In this section we indicate how the estimates $\widehat{\beta}$, $\widehat{m}$, $\overline{m}$ and the test statistic can be numerically computed. The algorithm can be motivated as follows. Consider the parametric (profile) likelihood function

$$\mathcal{L}(\beta) = \sum_{i=1}^{n} L(\mu_{i,\beta}, y_i), \tag{12}$$

$\mu_{i,\beta} = G\{x_i^T\beta + m_\beta(t_i)\}$. This function is optimized to obtain an estimate for β. The *smoothed* or *local* likelihood

$$\mathcal{L}^h\left(m_\beta(t)\right) = \sum_{i=1}^{n} \mathcal{K}_h(t - t_i)\, L\{\mu_{i,m_\beta(t)}, y_i\}, \tag{13}$$

$\mu_{i,m_\beta(t)} = G\{x_i^T\beta + m_\beta(t)\}$ is optimized to estimate the smooth function $m_\beta(t)$ at point t. The local weights $\mathcal{K}_h(t - t_i)$ here denote kernel weights with $\mathcal{K}$ denoting a kernel function and h the bandwidth.

Abbreviate now $m_j = m_\beta(t_j)$ and the individual log–likelihood in y_i by

$$\ell_i(\eta) = L\{G(\eta), y_i\}.$$

In the following, ℓ_i' and ℓ_i'' denote the derivatives of $\ell_i(\eta)$ with respect to η. The maximization of the local likelihood (13) requires to solve

$$0 = \sum_{i=1}^{n} \ell_i'(x_i^T\beta + m_j)\, \mathcal{K}_h(t_i - t_j). \tag{14}$$

For β we have from (12) to solve

$$0 = \sum_{i=1}^{n} \ell_i'(x_i^T\beta + m_i)\, \{x_i + m_i'\}. \tag{15}$$

A further differentiation of (14) leads to an expression for the derivative m_j' of m_j with respect to β

$$m_j' = -\frac{\sum_{i=1}^{n} \ell_i''(x_i^T\beta + m_j)\, \mathcal{K}_h(t_i - t_j) x_i}{\sum_{i=1}^{n} \ell_i''(x_i^T\beta + m_j)\, \mathcal{K}_h(t_i - t_j)}. \tag{16}$$

Equations (14)–(15) imply the following iterative Newton–Raphson type algorithm. Alternatively, the functions ℓ_i'' can be replaced by their expectations (w.r.t. to y_i) to obtain a Fisher scoring type procedure.

Profile Likelihood Algorithm

- *updating step for β*

$$\beta^{new} = \beta - \mathcal{B}^{-1} \sum_{i=1}^{n} \ell_i'(x_i^T \beta + m_i)\, \widetilde{x}_i$$

with a Hessian type matrix

$$\mathcal{B} = \sum_{i=1}^{n} \ell_i''(x_i^T \beta + m_i)\, \widetilde{x}_i \widetilde{x}_i^T$$

and

$$\widetilde{x}_j = x_j - \frac{\sum_{i=1}^{n} \ell_i''(x_i^T \beta + m_j)\, \mathcal{K}_h(t_i - t_j)\, x_i}{\sum_{i=1}^{n} \ell_i''(x_i^T \beta + m_j)\, \mathcal{K}_h(t_i - t_j)}.$$

- *updating step for m_j*

$$m_j^{new} = m_j - \frac{\sum_{i=1}^{n} \ell_i'(x_i^T \beta + m_j)\, \mathcal{K}_h(t_i - t_j)}{\sum_{i=1}^{n} \ell_i''(x_i^T \beta + m_j)\, \mathcal{K}_h(t_i - t_j)}.$$

The updating step for m_j is of quite complex structure. In some models (in particular for identity and exponential link functions G) equation (14) can be solved explicitly for m_j. For more details on this algortihm and possible simplifications we refer to Müller (1997).

To obtain the bias corrected parametric estimate $\overline{m}$, one has only to apply the updating step for $m_j = m_\beta(t_j)$, keeping $\widetilde{\beta}$ fixed.