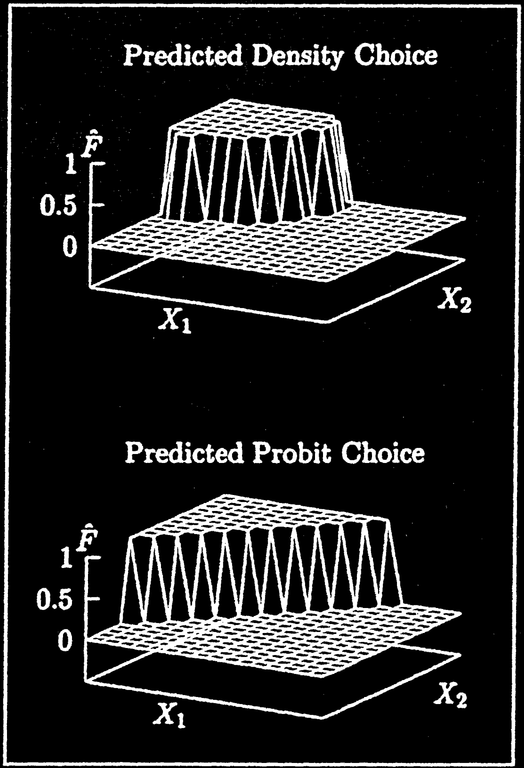
STATISTICS: textbooks and monographs volume 165

HANDBOOK OF APPLIED ECONOMETRICS AND STATISTICAL INFERENCE



EDITED BY
AMAN ULLAH
ALAN T. K. WAN
ANOOP CHATURVEDI

28

Semiparametric Panel Data Estimation: An Application to Immigrants' Homelink Effect on U.S. Producer Trade Flows

AMAN ULLAH University of California, Riverside, Riverside, California

KUSUM MUNDRA San Diego State University, San Diego, California

1. INTRODUCTION

Panel data refers to data where we have observations on the same cross-section unit over multiple periods of time. An important aspect of the panel data econometric analysis is that it allows for cross-section and/or time heterogeneity. Within this framework two types of models are mostly estimated; one is the fixed effect (FE) and the other is the random effect. There is no agreement in the literature as to which one should be used in empirical work; see Maddala (1987) for a good discussion on this subject. For both types of models there is an extensive econometric literature dealing with the estimation of linear parametric models, although some recent works on nonlinear and latent variable models have appeared; see Hsiao (1985), Baltagi (1998), and Mátyás and Sevestre (1996). It is, however, well known that the parametric estimators of linear or nonlinear models may become inconsistent if the model is misspecified. With this in view, in this paper we consider only the FE panel models and propose semiparametric estimators which are robust to the misspecification of the functional forms.

The asymptotic properties of the semiparametric estimators are also established.

An important objective of this paper is to explore the application of the proposed semiparametric estimator to study the effect of immigrants' "home link" hypothesis on the U.S. bilateral trade flows. The idea behind the home link is that when the migrants move to the U.S. they maintain ties with their home countries, which help in reducing transaction costs of trade through better trade negotiations, hence effecting trade positively. In an important recent work, Gould (1994) analyzed the home link hypothesis by considering the well-known gravity equation (Anderson 1979, Bergstrand 1985) in the empirical trade literature which relates the trade flows between two countries with economic factors, one of them being transaction cost. Gould specifies the gravity equation to be linear in all factors except transaction cost, which is assumed to be a nonlinear decreasing function of the immigrant stock in order to capture the home link hypothesis.* The usefulness of our proposed semiparametric estimators stems from the fact that the nonlinear functional form used by Gould (1994) is misspecified, as indicated in Section 3 of this paper. Our findings indicate that the immigrant home link hypothesis holds for producer imports but does not hold for producer exports in the U.S. between 1972 and 1980.

The plan of this paper is as follows. In Section 2 we present the FE model and proposed semiparametric estimators. These semiparametric estimators are then used to analyze the "home link" hypothesis in Section 3. Finally, the Appendix discusses the aysmptotic properties of the semiparametric estimators.

2. THE MODEL AND ESTIMATORS

Let us consider the parametric FE model as

$$y_{it} = x_{it}'\beta + z_{it}'\gamma + \alpha_i + u_{it} \qquad (i = 1, \cdots, n;\ t = 1, \cdots, T) \tag{2.1}$$

where y_{it} is the dependent variable, x_{it} and z_{it} are the $p \times 1$ and $q \times 1$ vectors, respectively, β, γ, and α_i are the unknown parameters, and u_{it} is the random error with $E(u_{it} \mid x_{it}, z_{it}) = 0$. We consider the usual panel data case of large n and small T. Hence all the asymptotics in this paper are for $n \to \infty$ for a fixed value of T. Thus, as $n \to \infty$, $\sqrt{nT}$ consistency and $\sqrt{n}$ consistency are equivalent.

*Transaction costs for obtaining foreign market information about country j in the U.S. used by Gould (1994) in his study is given by $A_e^{-\rho[M_{US,j}/(\theta+MU_{US,j})]}$, $\rho > 0$, $\theta > 0$, $A > 0$, where $M_{US,j}$ = stock of immigrants from country j in the United States.

From (2.1) we can write

$$Y_{it} = X'_{it}\beta + Z'_{it}\gamma + U_{it} \tag{2.2}$$

where $R_{it} = r_{it} - \bar{r}_{i.}$, $\bar{r}_{i.} = \sum_t^T r_{it}/T$. Then the well-known parametric FE estimators of β and γ are obtained by minimizing $\sum_i \sum_t U_{it}^2$ with respect to β and γ or $\sum_i \sum_t u_{it}^2$ with respect to β, γ, and α_i. These are the consistent least-squares (LS) estimators and are given by

$$\begin{aligned} b_p &= \left[\sum_i \sum_t (X_{it} - \tilde{X}_{it})(X_{it} - \tilde{X}_{it})'\right]^{-1} \sum_i \sum_t (X_{it} - \tilde{X}_{it}) Y_{it} \\ &= \left(S_{X-\tilde{X}}\right)^{-1} S_{X-\tilde{X},Y} \\ &= \left(X'M_Z X\right)^{-1} X'M_{ZY} \end{aligned} \tag{2.3}$$

and

$$c_p = S_Z^{-1}\left(S_{Z,Y} - S_X b_p\right) \tag{2.4}$$

where p represents parametric, $\tilde{X}'_{it} = Z'_{it}(\sum_i \sum_t Z_{it} Z'_{it})^{-1} \sum_i \sum_t Z_{it} X'_{it}$, $S_{A,B} = A'B/_{nT} = \sum_i \sum_t A_{it} B'_{it}/_{nT}$ for any scalar or column vector sequences A_{it} and B_{it}, $S_A = S_{A,A}$, and $M_Z = I - Z(Z'Z)^{-1}Z'$. The estimator $\hat{\alpha}_i = \bar{y}_{i.} - \bar{x}'_{i.}b_p - \bar{z}'_{i.}c_p$ is not consistent, and this will also be the case with the semiparametric estimators given below.

New semiparametric estimators of β and γ can be obtained as follows. From (2.2) let us write

$$E(Y_{it}|Z_{it}) = E\left(X'_{it}|Z_{it}\right)\beta + Z'_{it}\gamma \tag{2.5}$$

Then, subtracting (2.5) from (2.2), we get

$$Y^*_{it} = X^{*\prime}_{it}\beta + U_{it} \tag{2.6}$$

which gives the LS estimator of β as

$$\hat{\beta}_{sp} = \left(\sum_i \sum_t X^*_{it} X^{*\prime}_{it}\right)^{-1} \sum_i \sum_t X^*_{it} Y^*_{it} = S^{-1}_{X^*} S_{X^*,Y^*} \tag{2.7}$$

where $R^*_{it} = R_{it} - E(R_{it}|Z_{it})$ and sp represents semiparametric. We refer to this estimator as the semiparametric estimator, for the reasons given below.

The estimator $\hat{\beta}_{sp}$ is not operational since it depends on the unknown conditional expectations $E(A_{it}|Z_{it})$, where A_{it} is Y_{it} or X_{it}. Following Robinson (1988), these can however be estimated by the nonparametric kernel estimators

$$\hat{A}_{it} = \sum_j \sum_s A_{js} K_{it,js} \Big/ \sum_j \sum_s K_{it,js} \tag{2.8}$$

where $K_{it,js} = K((Z_{it} - Z_{js})/a)$, $j = 1, \ldots, n$; $s = 1, \ldots, T$, is the kernel function and a is the window width. We use product kernel $K(Z_{it}) = \prod_{l=1}^{q} k(Z_{it}, l)$, k is the univariate kernel and $Z_{it,l}$ is the ℓth component of Z_{it}. Replacing the unknown conditional expectations in (2.7) by the kernel estimators in (2.8), an operational version of $\hat{\beta}_{sp}$ becomes

$$b_{sp} = \left(\sum_i \sum_t (X_{it} - \hat{X}_{it})(X_{it} - \hat{X}_{it})'\right)^{-1} \sum_i \sum_t (X_{it} - \hat{X}_{it})(Y_{it} - \hat{Y}_{it})$$

$$= S^{-1}_{X-\hat{X}} S_{X-\hat{X}, Y-\hat{Y}} \tag{2.9}$$

Since the unknown conditional expectations have been replaced by their nonparametric estimates we refer to b_{sp} as the semiparametric estimator. After we get b_{sp},

$$c_{sp} = S_Z^{-1}(S_{Z,Y} - S_{Z,X}\, b_{sp}) \tag{2.10}$$

The consistency and asymptotic normality of b_{sp} and c_{sp} are discussed in the Appendix.

In a special case where we assume the linear parametric form of the conditional expectation, say $E(A_{it}|Z_{it}) = Z'_{it}\delta$, we can obtain the LS predictor as $\tilde{A}_{it} = Z'_{it}(\sum_i \sum_t Z_{it} Z'_{it})^{-1} \sum_i \sum_t Z_{it} A_{it}$. Using this in (2.7) will give $\hat{\beta}_{sp} = b_p$. It is in this sense that b_{sp} is a generalization of b_p for situations where, for example, X and Z have a nonlinear relationship of unknown form.

Both the parametric estimators b_p, c_p and the semiparametric estimators b_{sp}, c_{sp} described above are the $\sqrt{n}$ consistent global estimators in the sense that the model (2.2) is fitted to the entire data set. Local pointwise estimators of β and γ can be obtained by minimizing the kernel weighted sum of squares

$$\sum_i \sum_t [y_{it} - x'_{it}\beta - z'_{it}\gamma - \alpha_i]^2 K\left(\frac{x_{it} - x}{h}, \frac{z_{it} - z}{h}\right) \tag{2.11}$$

with respect to β, γ, and α ; h is the window width. The local pointwise estimators so obtained can be denoted by $b_{sp}(x, z)$ and $c_{sp}(x, z)$, and these are obtained by fitting the parametric model (2.1) to the data close to the points x, z, as determined by the weights $K()$. These estimators are useful for studying the local pointwise behaviors of β and γ, and their expressions are given by

$$\begin{bmatrix} b_{sp}(x,z) \\ c_{sp}(x,z) \end{bmatrix} = \left(\sum_i \sum_t (w_{it} - \hat{w}_{i.})(w_{it} - \hat{w}_{i.})' \right)^{-1} \sum_i \sum_t (w_{it} - \hat{w}_{i.})(y_{it.} - \hat{y}_{i.})$$
$$= S^{-1}_{w-\hat{w}.} S_{w-\hat{w},y-\hat{y}} \tag{2.12}$$

where $w'_{it} = [\sqrt{K_{it}}x'_{it}\sqrt{K_{it}}z'_{it}]$, $K_{it} = K((x_{it} - x)/h, (z_{it} - z)/h)$, $\hat{A}_i = \sum_t A_{it}K_{it}/\sum_t K_{it}$.

While the estimators b_p, c_p and b_{sp}, c_{sp} are the $\sqrt{n}$ consistent global estimators, the estimators $b_{sp}(x,z)$, $c_{sp}(x,z)$ are the $\sqrt{nh^{p+q+2}}$ consistent local estimators (see Appendix). These estimators also provide a consistent estimator of the semiparametric FE model

$$y_{it} = m(x_{it}, z_{it}) + \alpha_i + u_{it} \tag{2.13}$$

where $m()$ is the nonparametric regression. This model is semiparametric because of the presence of the parameters α_i. It is indicated in the Appendix that

$$\hat{m}_{sp}(x_{it}, z_{it}) = x'_{it}b_{sp}(x_{it}, z_{it}) + z'_{it}c_{sp}(x_{it}, z_{it}) \tag{2.14}$$

is a consistent estimator of the unknown function $m(x_{it}, z_{it})$, and hence b_{sp}, c_{sp} are the consistent estimators of its derivatives. In this sense $\hat{m}_{sp}(x_{it}, z_{it})$ is a local linear nonparametric regression estimator which estimates the linear model (2.1) nonparametrically; see Fan (1992, 1993) and Gozalo and Linton (1994). We note however the well-known fact that the parametric estimator $x'_{it}b_p + z'_{it}c_p$ is a consistent estimator only if $m(x_{it}, z_{it}) = x'_{it}\beta + z'_{it}\gamma$ is the true model. The same holds for any nonlinear parametric specification estimated by the global parametric method, such as nonlinear least squares.

In some situations, especially when the model (2.13) is partially linear in x but nonlinear of unknown form in z, as in Robinson (1988), we can estimate β globally but γ locally and vice-versa. In these situations we can first obtain the global $\sqrt{n}$ consistent estimate of β by b_{sp} in (2.9). After this we can write

$$y^o_{it} = y_{it} - x'_{it}b_{sp} = z'_{it}\gamma + \alpha_i + v_{it} \tag{2.15}$$

where $v_{it} = u_{it} + x'_{it}(\beta - b_{sp})$. Then the local estimation of γ can be obtained by minimizing

$$\sum_i \sum_t [y^o_{it} - z'_{it}\gamma - \alpha_i]^2 K\left(\frac{z_{it} - z}{h}\right) \tag{2.16}$$

which gives

$$c_{sp}(z) = \left(\sum_i \sum_t (z_{it} - \hat{z}_{i.})(z_{it} - \hat{z}_{i.})' K_{it}\right)^{-1} \sum_i \sum_t (z_{it} - \hat{z}_{i.})(y^o_{it} - \hat{y}^o_{it.}) K_{it}$$

$$= S^{-1}_{\sqrt{K}(z-\hat{z}.)} S_{\sqrt{K}(z-\hat{z}.)(y^o - \hat{y}.^o)} \tag{2.17}$$

where $K_{it} = K((z_{it} - z)/h)$ and $\hat{A}_i = \sum_t A_{it} K_{it} / \sum_t K_{it}$. Further, $\hat{\alpha}_i(z) = \hat{y}^o_{i.} - \hat{z}_{i.}\, c_{sp}(z)$. As in (2.14), $\hat{m}_{sp}(z_{it}) = z'_{it}\, c_{sp}(z)$ is a consistent local linear estimator of the unknown nonparametric regression in the model $y^o_{it} = m(z_{it}) + \alpha_i + u_{it}$. But the parametric estimator $z'_{it}\, \hat{\gamma}_p$ will be consistent only if $m(z_{it}) = z'_{it}\, \gamma$ is true. For discussion on the consistency and asymptotic normality of $b_{sp}(z)$, $c_{sp}(z)$, and $\hat{m}_{sp}(z)$, see Appendix.

3. MONTE CARLO RESULTS

In this section we discuss Monte Carlo simulations to examine the small sample properties of the estimator given by (2.9). We use the following data generating process (DGP):

$$y_{it} = \beta x_{it} + \delta z_{it} + \gamma z^2_{it} + u_{it}$$
$$= \beta x_{it} + m(z_{it}) + u_{it} \tag{3.1}$$

where z_{it} is independent and uniformly distributed in the interval $[-\sqrt{3}, \sqrt{3}]$, x_{it} is independent and uniformly distributed in the interval $[-\sqrt{5}, \sqrt{5}]$, u_{it} is i.i.d. $N(0, 5)$. We choose $\beta = 0.7$, $\delta = 1$; and $\gamma = 0.5$. We report estimated bias, standard deviation (Std) and root mean squares errors (Rmse) for the estimators. These are computed via Bias $(\hat{\beta}) = M^{-1} \sum^M (\hat{\beta}_j - \beta_j)$, Std$(\hat{\beta}) = \{M^{-1} \sum^M (\hat{\beta}_j - \text{Mean}(\hat{\beta}))^2\}^{1/2}$, and Rmse$(\hat{\beta}) = \{M^{-1} \sum^M (\hat{\beta}_j - \beta)^2\}^{1/2}$, where $\hat{\beta} = b_{sp}$, M is the number of replications and $\hat{\beta}_j$ is the jth replication. We use $M = 2000$ in all the simulations. We choose $T = 6$ and $n = 50, 100, 200$, and 500. The simulation results are given in Table 1. The results are not dependent on δ and γ, so one can say that the results are not sensitive to different functional forms of $m(z_{it})$. We see that Std and Rmse are falling as n increases.

4. EMPIRICAL RESULTS

Here we present an empirical application of our proposed semiparametric estimators. In this application we look into the effect of the immigrants' "home link" hypothesis on U.S. bilateral producer trade flows. Immigration

Table 1. The case of $\beta = 0.7$, $\delta = 1$, $\gamma = 0.5$

	b_{sp}		
	Bias	Std	Rmse
$n = 50$	−0.116	0.098	0.152
$n = 100$	−0.115	0.069	0.134
$n = 200$	−0.118	0.049	0.128
$n = 500$	−0.117	0.031	0.121

has been an important economic phenomenon for the U.S., with immigrants varying in their origin and magnitude. A crucial force in this home link is that when migrants move to the U.S. they maintain ties with their home countries, which helps in reducing transaction costs of trade through better trade negotiations, removing communication barriers, etc. Migrants also have a preference for home products, which should effect U.S. imports positively. There have been studies to show geographical concentrations of particular country-specific immigrants in the U.S. actively participating in entrepreneurial activities (Light and Bonacich 1988). This is an interesting look at the effect of immigration other than the effect on the labor market, or welfare impacts, and might have strong policy implications for supporting migration into the U.S. from one country over another.

A parametric empirical analysis of the "home link" hypothesis was first done by Gould (1994). His analysis is based on the gravity equation (Anderson 1979, Bergstrand 1985) extensively used in the empirical trade literature, and it relates trade flows between two countries with economic forces, one of them being the transaction cost. Gould's important contribution specifies the transaction cost factor as a nonlinear decreasing function of the immigrant stock to capture the home link hypothesis: decreasing at an increasing rate. Because of this functional form the gravity equation becomes a nonlinear model, which he estimates by nonlinear least squares using an unbalanced panel of 47 U.S. trading partners.

We construct a balance panel of 47 U.S. trading partners over nine years (1972–1980), so here $i = 1, \ldots, 47$ and $t = 1, \ldots, 9$, giving 423 observations. The country specific effects on heterogeneity are captured by the fixed effect. In our case, y_{it} = manufactured U.S. producers' exports and imports, x_{it} includes lagged value of producers' exports and imports, U.S. population, home-country population, U.S. GDP, home-country GDP, U.S. GDP deflator, home-country GDP deflator, U.S. export value index, home-country export value index, U.S. import value index, home-country import value index, immigrant stay, skilled–unskilled ratio of the migrants, and z_{it} is immigrant stock to the U.S. Data on producer-manufactured imports and exports were taken from OECD statistics. Immigrant stock, skill level and length of stay of migrants were taken from INS public-use data on yearly immigration. Data on income, prices, and population were taken from IMF's International Financial Statistics.

We start the analysis by first estimating the immigrants' effect on U.S. producer exports and imports using Gould's (1994) parametric functional form and plot it together with the kernel estimation; see Figures 1 and 2. The kernel estimator is based on the normal kernel given as $K((z_{it} - z)/h) = 1/\sqrt{2\pi}\exp\{-(1/2)((z_{it} - z)/h)^2\}$ and h, the window-width, is taken as $cs(nT)^{-1/5}$, c is a constant, and s is the standard derivation for variable z; for details on the choice of h and K see Härdle (1990) and Pagan and Ullah (1999). Comparing the results with the actual trade flows, we see from Figures 1 and 2 that the functional form assumed in the parametric estima-

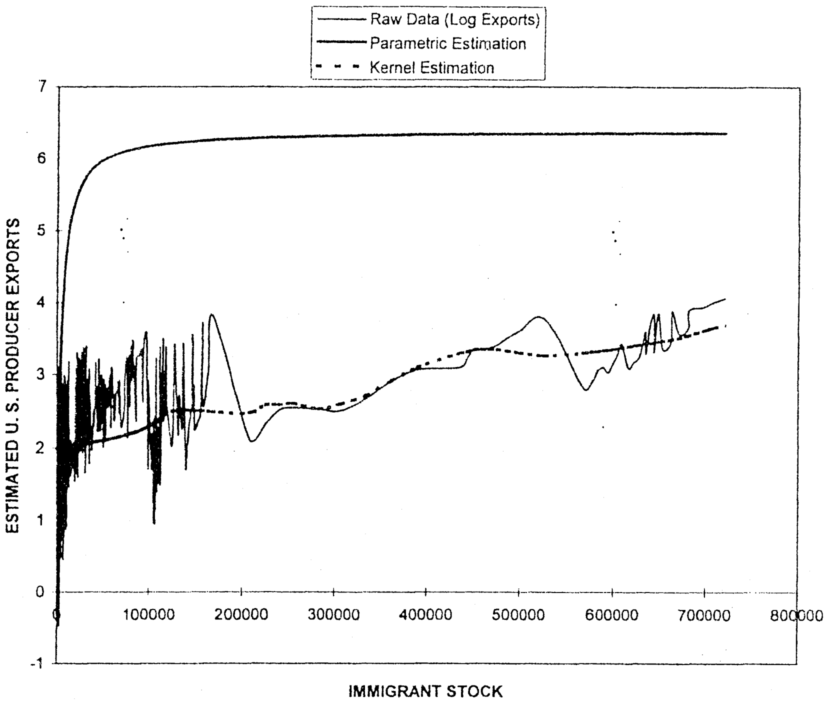


Figure 1. Comparison of U.S. producer exports with parametric functional estimation and kernel estimation.

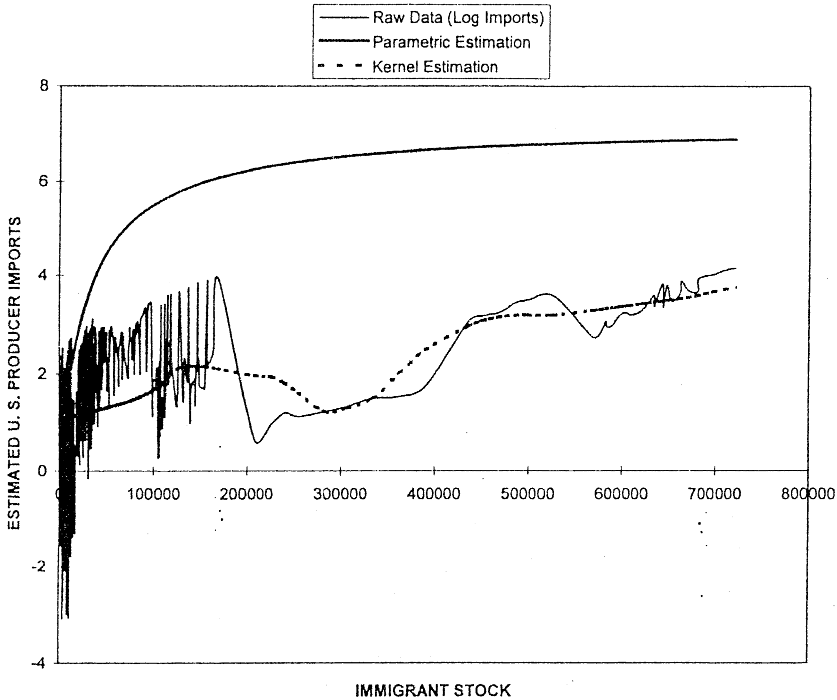


Figure 2. Comparison of U.S. producer imports with parametric functional estimation and kernel estimation

tion is incorrect and hence Gould's nonlinear LS estimates may be inconsistent. In fact the parametric estimates, b_p and c_p, will also be inconsistent. In view of this we use our proposed $\sqrt{n}$ consistent semiparametric estimator of β, b_{sp}, in (2.9) and the consistent semiparametric local linear estimator of γ, $c_{sp}(z)$, in (2.17).

First we look at the semiparametric estimates b_{sp} given in Table 2. The immigrant skilled–unskilled ratio affects exports and imports positively, though it is insignificant. This shows that skilled migrants are bringing better foreign market information. As the number of years the immigrant stays in the U.S. increases, producer exports and producer imports fall at an increasing rate. It can be argued that the migrants change the demand structure of the home country adversely, decreasing U.S. producer exports and supporting imports. But once the home country information, which they carry becomes obsolete and their tastes change, their effect on the trade falls. When the inflation index of a country rises, exports from that

Table 2. Bilateral manufactured producer trade flows between the U.S. and the immigrant home countries

	U.S. producer exports		U.S. producer imports	
Dependent variable	Parametric model	SPFE	Parametric model	SPFE
U.S. GDP deflator	0.52	−9.07	12.42	5.45
	(3.34)	(18.22)	(9.69)	(77.62)
Home-country GDP deflator	−0.25	−0.09	0.29	−0.11
	(0.09)[a]	(0.06)	(0.26)	(0.35)
U.S. GDP	−1.14	−3.29	6.71	5.35
	(2.13)	(11.01)	(6.71)	(53.74)
Home-country GDP	0.60	0.17	0.56	−0.16
	(0.11)[a]	(0.09)[c]	(0.34)[b]	(0.45)[a]
U.S. population	5.09	88.24	6.05	−67.18
	(40.04)	(236.66)	(123.8)	(1097.20)
Home-country population	0.41	0.58	0.58	−5.31
	(0.18)[a]	(0.48)	(0.53)[c]	(2.47)
Immigrant stay	−0.06	0.01	−0.16	−0.13
	(0.05)	(0.25)	(0.01)	(1.18)
Immigrant stay (squared)	0.002	0.001	0.01	0.003
	(0.003)	(0.02)	(0.01)	(0.07)
Immigrant skilled–unskilled ratio	0.01	0.02	0.06	0.02
	(0.02)	(0.02)	(0.06)	(0.06)
U.S. export unit value index	1.61	1.91		
	(0.46)[a]	(0.57)[a]		
Home-country import unit value index	−0.101	0.072		
	(0.04)	(0.09)		
Home-country export unit value index			1.72	0.37
			(0.77)[a]	(1.85)
U.S. import unit value index			−0.10	0.004
			(0.34)	(0.22)

Newey–West corrected standard errors in parentheses. [a]Significant at 1% level. [b]Significant at 5% level. [c]Significant at 10% level.

country may become expensive and are substituted by domestic production in the importing country. Hence, when the home-country GDP deflator is going up, U.S. producer imports fall and the U.S. GDP deflator affects U.S. producer exports negatively. The U.S. GDP deflator has a positive effect on U.S. imports, which might be due to the elasticity of substitution among imports exceeding the overall elasticity between imports and domestic production in the manufactured production sector in the U.S., whereas the opposite holds in the migrants' home country. The U.S. export value index reflects the competitiveness for U.S. exports and has a significant positive effect on producer exports. This may be due to the supply elasticity of transformation among U.S. exports exceeding the overall elasticity between exports and domestic goods, which is true for the home-country export unit value index too. The U.S. and the home country import unit value indexes have a positive effect on producer imports and producer exports respectively. This shows that the elasticity of substitution among imports exceeds the overall elasticity between domestic and imported goods, both in the U.S. and in the home country. The immigrants' home-country GDP affects the producer exports positively and is significant at the 10% level of significance. The U.S. GDP affects producer exports negatively and also the home-country GDP affects producer imports negatively, showing that the demand elasticity of substitution among imports is less than unity both for the U.S. and its trading partners.

To analyze the immigrant "home link" hypothesis, which is an important objective here, we obtain elasticity estimates $c_{sp}(z)$ at different immigrant stock levels for both producer's exports and producer's imports. This shows how much U.S. bilateral trade with the ith country is brought about by an additional immigrant from that country. Based on this, we also calculate in Table 3 the average dollar value change (averaged over nine years) in U.S. bilateral trade flows: $\bar{c}_{isp} \times \bar{z}_i$, where $\bar{c}_{isp} = \sum_t c_{sp}(z_{it})/T$ and $\bar{z}_i = \sum_t z_{it}/T$ is the average immigrant stock into the U.S. from the ith country. When these values are presented in Figures 3 and 4, we can clearly see that the immigrant home link hypothesis supports immigrant stock affecting trade positively for U.S. producer imports but not for U.S. producer exports. These findings suggest that immigrant stock and U.S. producer imports are complements in general, but the immigrants and producer exports are substitutes. In contrast, Gould's (1994) nonlinear parametric framework suggests support for the migrants' "homelink hypothesis" for both exports and imports. The difference in our results for exports with those of Gould may be due to misspecification of the nonlinear transaction cost function in Gould and the fact that he uses unbalanced panel data. All these results however indicate that the "home link" hypothesis alone may not be sufficient to look at the broader effect of immigrant stock on bilateral trade flows. The labor role of migrants and the welfare effects of immigration, both in the receiving and the sending country, need to be taken into account. These results also crucially depend on the sample period; during the 1970s the U.S. was facing huge current account deficits. In any case, the above analysis does open interesting questions as to what should be the U.S. policy on immigration; for example, should it support more immigration from one country over another on the basis of dollar value changes in import or export?

Table 3. Average dollar value change in U.S. producer trade flows from one additional immigrant between 1972 and 1980

Country		Producer exports	Producer imports
1	Australia	−84 447.2	107 852.2
2	Austria	−257 216	332 576.7
3	Brazil	−72 299.9	91 995.54
4	Canada	−1 908 566	2 462 421
5	Colombia	−300 297	381 830.7
6	Cyprus	−11 967.4	15 056.1
7	Denmark	−65 996.3	85 321.2
8	El Salvador	−115 355	146 500.3
9	Ethiopia	−11 396.6	13 098.77
10	Finland	−93 889.6	121 071.7
11	France	−174 535	225 599.7
12	Greece	−557 482	718 292.1
13	Hungary	−172 638	163 015.4
14	Iceland	−13 206.8	17 003.16
15	India	−311 896	383 391.8
16	Ireland	−577 387	742 629.5
17	Israel	−126 694	159 101.8
18	Italy	−2 356 589	3 045 433
19	Japan	−446 486	575 985.8
20	Jordan	−33 074.7	41 427
21	Kenya	−3 604.1	4 044.627
22	Malaysia	−9 761.78	11 766
23	Malta	−23 507.1	30 184.8
24	Morocco	−2 899.56	2 797.519
25	Netherlands	−346 098	447 181.1
26	New Zealand	−23 666.3	30 182.7
27	Nicaragua	−74 061.1	93 930.9
28	Norway	−231 098	298 533.2
29	Pakistan	−35 508.4	42 682.64
30	Philippines	−214 906	258 027.4
31	S. Africa	−29243.3	37247.1
32	S. Korea	−89567.5	109286.9
33	Singapore	−4095.1	4863.85
34	Spain	−161804	207276.4
35	Sri Lanka	−7819.8	9685.5
36	Sweden	−220653	28500.9
37	Switzerland	−91599.2	118259.2

Table 3. Continued

Country		Producer exports	Producer imports
38	Syria	−358 830.3	44 644.6
39	Tanzania	−2 875.3	2 679.2
40	Thailand	−49 734.8	58 071.3
41	Trinidad	−113 210	142 938.1
42	Tunisia	−3 285.2	3 066.1
43	Turkey	−115 192	147 409.5
44	U.K.	0	0
45	W. Germany	−193 8678	2505652
46	Yugoslavia	−468 268	598 664.1
47	Zimbabwe	−2 209.5	1 997.1

ACKNOWLEDGMENTS

The first author gratefully acknowledges the research support from the Academic Senate, UCR. The authors are thankful for useful discussions with Q. Li and the comments by the participants of the Western Economics Association meeting at Seattle. They are also grateful to D. M. Gould for providing the data used in this paper.

APPENDIX

Here we present the asymptotic properties of the estimators in Section 2. First we note the well-known results that, as $n \to \infty$,

$$\left.\begin{aligned} \sqrt{nT}(b_p - \beta) &\sim N\Big(0, \sigma^2 \big(P \lim S_{X-\hat{X}}\big)^{-1}\Big) \\ \sqrt{nT}(c_p - \beta) &\sim N\Big(0, \sigma^2 \big(P \lim S_{Z-\tilde{Z}}\big)^{-1}\Big) \end{aligned}\right\} \tag{A.1}$$

where $\tilde{Z}$ is generated by $\tilde{Z}'_{it} = X'_{it}(\sum_i \sum_t X_{it} X'_{it})^{-1} \sum_i \sum_t X_{it} Z'_{it}$ and $P \lim$ represents probability limit; see the book White (1984).

Next we describe the assumptions that are needed for the consistency and asymptotic normality of b_{sp}, c_{sp}, $b_{sp}(x, z)$, $c_{sp}(x, z)$, and $c_{sp}(z)$ given above. Following Robinson (1988), let G^{λ}_{μ} denote the class of functions such that if $g \in G^{\lambda}_{\mu}$, then g is μ times differentiable; g and its derivatives (up to order μ) are all bounded by some function that has λth-order finite moments. Also,

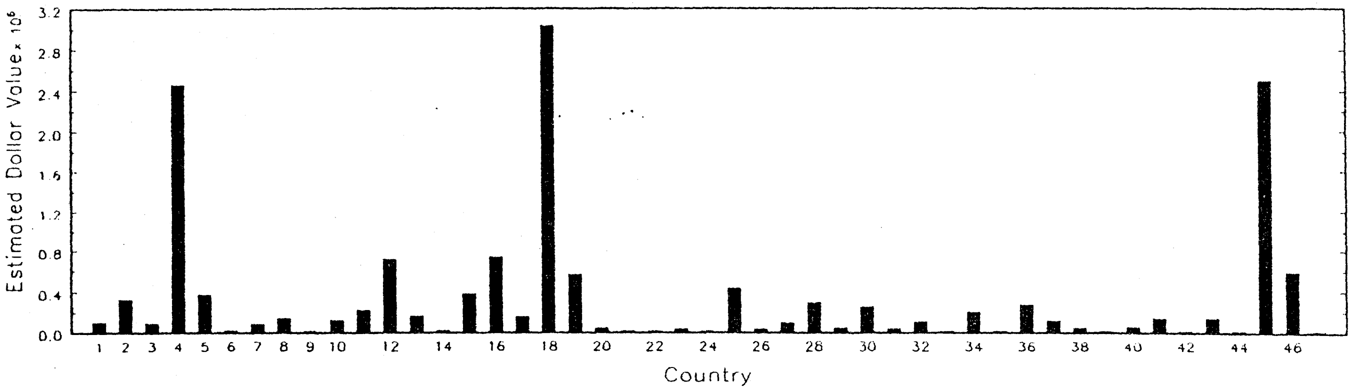


Figure 3. Average estimated dollar value change in U.S. producer imports from one additional immigrant.

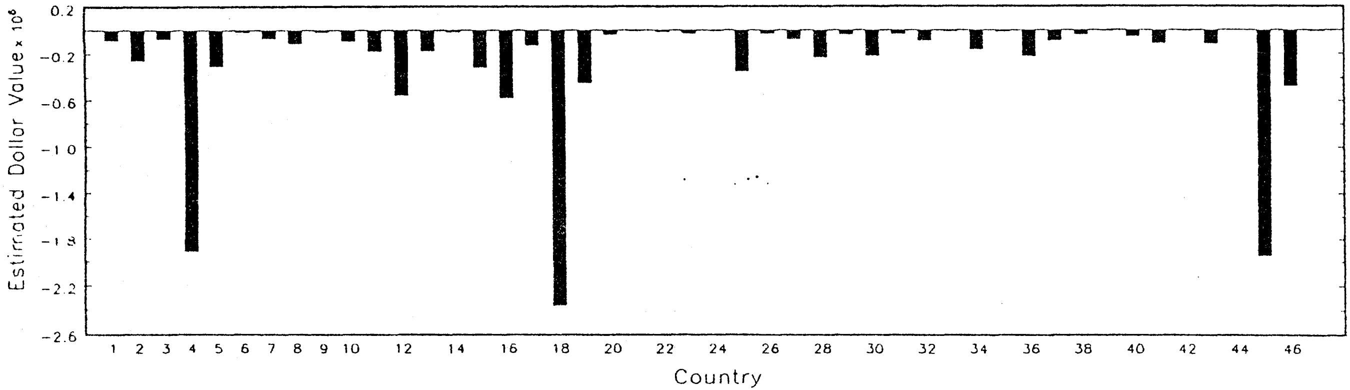


Figure 4. Average estimated dollar value change in the U.S. producer exports from one additional immigrant.

K_2 denotes the class of nonnegative kernel functions k: satisfying $\int k(v)v^m dv = \delta om$ for $m = 0, 1$ (δom is the Kronecker's delta), $\int k(v)vv'dv = C_k I$ $(I > 0)$, and $k(u) = O((1 + |u|^{3+\eta})^{-1})$ for some $\eta > 0$. Further, we denote $\int k^2(v)vv'd$ $v = D_k.I$. We now state the following assumptions:

(A1) (*i*) for all $t, (y_{it}, x_{it}, z_{it})$ are i.i.d. across i and z_{it} admits a density function $f \in G^{\infty}_{\mu-1}$, $E(x|z)$ and $E(z|x) \in G^4_\mu$ for some positive integer $\mu > 2$; (ii) $E(u_{it}|x_{it}, z_{it}) = 0$, $E(u^2_{it}|x_{it}, z_{it}) = \sigma^2(x_{it}, z_{it})$ is continuous in x_{it} and z_{it}, and u_{it}, $\eta_{it} = x_{it} - E(x_{it}|z_{it})$, $\xi_{it} = (z_{it} - E(z_{it}|x_{it})$ have finite $(4+\delta)$th moment for some $\delta > 0$.

(A2) $\bar{k} \in K_\lambda$; as $n \to \infty$, $a \to 0$, $na^{4\lambda} \to 0$, and $na^{\max(2q-4,\, q)} \to \infty$.

(A3) $k \in K_2$ and $k(v) \geq 0$; as $n \to \infty$, $h \to 0$, $nh^{q+2} \to \infty$, and $nh^{q+4} \to 0$.

(A1) requires independent observations across i, and gives some moment and smoothness conditions. The condition (A2) ensures b_{sp} and c_{sp} are $\sqrt{n}$ consistent. Finally (A3) is used in the consistency and asymptotic normality of $b_{sp}(x, z)$; $c_{sp}(x, z)$, and $c_{sp}(z)$.

Under the assumptions (A1) and (A2), and taking $\sigma^2(x, z) = \sigma^2$ for simplicity, the asymptotic distributions of the semiparametric estimators b_{sp} and c_{sp} follow from Li and Stengos (1996), Li (1996) and Li and Ullah (1998). This is given by

$$\sqrt{nT}\left(b_{sp} - \beta\right) \sim N\left(0, \sigma^2\Sigma^{-1}\right) \text{ and } \sqrt{nT}\left(c_{sp} - \beta\right) \sim N\left(0, \sigma^2\Omega^{-1}\right) \qquad \text{(A.2)}$$

where $\Sigma = E(\eta'_1\eta_1/T)$ and $\Omega = E(\xi'_1\xi_1/T)$; $\eta'_i = (\eta_{i1}, \ldots, \eta_{iT})$. Consistent estimators for Σ^{-1} and Ω^{-1} are $\hat{\Sigma}^{-1}$ and $\hat{\Omega}^{-1}$, respectively, where $\hat{\Sigma} = (1/(\mathrm{nT}))\ \Sigma_i\Sigma_t(X_{it} - \hat{X}_{it})(X_{it} - \hat{X}_{it})' = 1/(nT))\Sigma_i(X_i - \hat{X}_i)'(X_i - \hat{X}_i)$ and $\hat{\Omega} = (1/(nT))\Sigma_i\Sigma_t(Z_{it} - \hat{Z}_{it})(Z_{it} - \hat{Z}_{it})'$.

The semiparametric estimators b_{sp} and c_{sp} depend upon the kernel estimators which may have a random denominator problem. This can be avoided by weighting (2.5) by the kernel density estimator $\hat{f}_{it} = \hat{f}(Z_{it}) = (1/(nTa^q))\sum_j\sum_s K_{it,js}$. This gives $\tilde{b}_{sp} = S^{-1}_{(X-\hat{X})\hat{f},(Y-\hat{Y})\hat{f}}$. In this case $\hat{\Sigma}$ will be the same as above with $X - \hat{X}$ replaced by $(X - \hat{X})\hat{f}$. Finally, under the assumptions (A1) to (A3) and noting that $(nTh^{q+2})^{1/2}(b_{sp} - \beta) = o_p(1)$, it follows from Kneisner and Li (1996) that for $n \to \infty$

$$\left(nTh^{q+2}\right)^{-1}\left(c_{sp(z)} - \gamma(z)\right) \sim N(0, \Sigma_1) \qquad \text{(A.3)}$$

where $\Sigma_1 = (\sigma^2(z)/f(z))C_k^{-1}D_kC_k^{-1}$, C_k and D_k are as defined above. In practice we replace $\sigma^2(z)$ by its consistent estimator $\hat{\sigma}^2(z_{it}) = \Sigma_j\Sigma_s(y^o_{js}$

$-z_{js}\, c_{sp}(z_{js}))^2 K_{it,js}/\Sigma_j \Sigma_s K_{it,js}$. Further, denoting $m(z) = z'\gamma$ and $\hat{m}_{sp}(z) = z' c_{sp}(z)$, as $n \to \infty$

$$(nTh^q)^{1/2}(\hat{m}_{sp}(z) - m(z)) \sim N(0, \Sigma_2) \tag{A.4}$$

where $\sum_2 = (O^2(z)/f(z)) \int K^2(v)dv$; see Gozalo and Linton (1994). Thus the asymptotic variance of $\hat{m}(z)$ is independent of the parametric model $z\gamma$ used to get the estimate $\hat{m}(z)$ and it is the same as the asymptotic variance of Fan's (1992, 1993) nonparametric local linear estimator. In this sense $c_{sp}(z)$ and $\hat{m}_{sp}(z)$ are the local linear estimators.

The asymptotic normality of the vector $[b'_{sp}(x, z), c'_{sp}(x, z)]$ is the same as the result in (A.3) with $q+2$ replaced by $p+q+2$ and z replaced by (x, z). As there, these estimators are also the local linear estimators.

REFERENCES

Anderson, J. E. (1979), A theoretical foundation for the gravity equation. *American Economics Review* 69: 106–116.

Baltagi, B. H. (1998), Panel data methods. *Handbook of Applied Economics Statistics* (A. Ullah and D. E. A. Giles, eds.)

Bergstrand, J. H. (1985), The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics* 67: 474–481.

Fan, J. (1992), Design adaptive nonparametric regression. *Journal of the American Statistical Association* 87: 998–1024.

Fan, J. (1993), Local linear regression smoothers and their minimax efficiencies. *Annals of Statistics* 21: 196–216.

Gould, D. (1994), Immigrant links to the home-country: Empirical implications for U.S. bilateral trade flows. *The Review of Economics and Statistics* 76: 302–316.

Gozalo, P., and O. Linton (1994), Local nonlinear least square estimation: using parametric information nonparametrically. *Cowler Foundation Discussion paper No. 1075.*

Härdle. W. (1990), *Applied Nonparametric Regression*, New York: Cambridge University Press.

Hsiao, C. (1985), Benefits and limitations of panel data. *Econometric Review* 4: 121–174.

Kneisner, T., and Q. Li (1996), Semiparametric panel data model with dynamic adjustment: Theoretical considerations and an application to labor supply. *Manuscript*, University of Guelph.

Li, Q. (1996), A note on root-*N*-consistent semiparametric estimation of partially linear models. *Economic Letters* 51: 277–285.

Li, Q., and T. Stengos (1996), Semiparametric estimation of partially linear panel data models. *Journal of Econometrics* 71: 389–397.

Li, Q., and A. Ullah (1998), Estimating partially linear panel data models with one-way error components errors. *Econometric Reviews* 17(2): 145–166.

Light, I., and E. Bonacich (1988), *Immigrant Entrepreneurs*. Berkeley: University of California Press.

Maddala, G. S. (1987), Recent developments in the econometrics of panel data analysis. *Transportation Research* 21: 303–326.

Mátyás, L., and P. Sevestre (1996), *The Econometrics of Panel Data: A Handbook of the Theory with Applications*. Dordrecht: Kluwer Academic Publishers.

Pagan, A., and A. Ullah (1999), *Nonparametric Econometrics*. New York: Cambridge University Press.

Robinson, P. M. (1988), Root-*N*-consistent semiparametric regression. *Econometrica* 56(4): 931–954.

White, H. (1984), *Asymptotic Theory for Econometricians*. New York: Academic Press.