



Assessing inherent model bias: An application to native displacement in response to immigration

Giovanni Peri ^{a,*}, Chad Sparber ^b

^a Department of Economics, UC Davis, One Shields Avenue, Davis, CA 95616, United States

^b Department of Economics, Colgate University, 13 Oak Drive, Hamilton, NY 13346, United States

ARTICLE INFO

Article history:

Received 3 February 2010

Revised 25 August 2010

Available online 24 September 2010

JEL classification:

J61

R23

Keywords:

Immigration
Microsimulation
Native displacement
Internal migration
Crowding out
Employment effects
Model bias

ABSTRACT

There is a long-standing debate among academics about the effect of immigration on native internal migration decisions. If immigrants displace natives this may indicate a direct cost of immigration in the form of decreased employment opportunity for native workers. Moreover, displacement would also imply that cross-region analyses of wage effects systematically underestimate the consequences of immigration. The widespread use of such area studies for the US and other countries makes it especially important to know whether a native internal response to immigration truly occurs. This paper introduces a microsimulation methodology to test for inherent bias in regression models that have been used in the literature. We show that some specifications have built biases into their models, thereby casting doubt on the validity of their results. We then provide a brief empirical analysis with a panel of observed US state-by-skill data. Together, our evidence argues against the existence of native displacement. This implies that cross-region analyses of immigration's effect on wages are still informative.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

There is a long-standing debate on whether immigration reduces the employment opportunities of natives. Economic analyses often exploit the wide variation in immigration rates across US states (or cities) and skill groups to identify whether immigration is associated with low native employment growth due to internal migration or job displacement across skill-state (or skill-city) cells. Though this correlation cannot definitively identify the effects of immigration (since causality is unclear and there may be omitted variables bias), researchers often cite such results as *prima facie* evidence for or against the crowding-out theory.

The importance of this issue is not limited to simply understanding the direct question of native displacement and employment opportunities. It also informs the validity of performing cross-regional analyses of the wage effects of immigration (or “area studies”). For example, most of the literature on US immigration across local labor markets finds little impact of immigration on wages.¹ These studies typically argue that mechanisms other than internal migration

allow each region to absorb the higher supply of workers.² If so, then cross-regional analysis is informative of the impact of immigrants on wages at the national level. In the presence of displacement, however, the wage effects of immigration would dissipate throughout the US – not just in the states receiving large numbers of immigrants. Thus, cross-region wage regressions would miss (or underestimate) the effects of immigration if displacement exists.³

In analyzing internal migration, researchers must make numerous methodological decisions. Should the unit of analysis be states, cities, or census tracts? Should regressions include a panel with fixed effects or employ just a single long-term cross-section? Which individuals should be included in the sample selection? Should regressions concern the population, labor force, or employees? These all are important questions to answer. This paper focuses on a most basic choice: how to specify the explanatory and dependent variables in the regression model that aims at estimating displacement. While this seems a trivial issue, we will show that some specifications in the literature may have built a bias into the estimates of the displacement coefficient they intended to identify.

* Corresponding author.

E-mail addresses: gperi@ucdavis.edu (G. Peri), csparber@mail.colgate.edu (C. Sparber).

¹ See Card (2001, 2007, 2009), Card and Lewis (2007), Ottaviano and Peri (2007) and Peri and Sparber (2009).

² Recent papers have proposed different mechanisms as margins of adjustment to immigration. Lewis (2005) indicates the choice of technique, Ottaviano and Peri (2008) focus on native-immigrant complementarities and capital adjustment, Peri and Sparber (2009) emphasize changes in relative specialization.

³ See Longhi et al. (2008) or Hanson (2008) for recent surveys.

We begin with a brief literature review in Section 2. It is far from exhaustive, but we focus on studies that have employed cross-regional internal migration regressions at similar levels of aggregation as this will facilitate model comparison. We highlight two seminal works: Card (2001) – which finds no evidence for displacement using US city data – and Borjas (2006) – which argues for large displacement effects at the city-level: roughly three natives are displaced for every 10 immigrants.

In Section 3 we ask whether the disparate conclusions in these and other studies can be a result of model specification. Our procedure is similar in spirit to Wolf (2001), who advocates using microsimulation to develop appropriate empirical models. We extend this idea by using microsimulation to test for inherent model bias. First we construct hypothetical data using data generating processes that assume, in turn, that the inflow of immigrants is negatively correlated, uncorrelated, or positively correlated with the inflow of natives. We then test whether previous empirical models are able to correctly identify the sign and magnitude of the underlying assumed correlation. Unfortunately, empirical model specification is not inconsequential. In particular, Borjas (2006) specifications are biased toward identifying displacement, and this bias grows larger as the variance of native flows rises in proportion to the variance of immigrant flows independently of their correlation.

This paper does not attempt to replicate the results of prior studies. Differences in sample selection, period of analysis, and other issues would encumber that endeavor while detracting from the main issue – we care to show the importance of model specification in correctly identifying the displacement effect of immigrants on natives. Nonetheless, Section 4 employs a number of alternative empirical specifications to briefly analyze the association between immigration and native migration using observed data from 32 skill-cells and 51 US states (including the District of Columbia) over Census years 1970–2000. Of the models we explore, only the Borjas (2006) specifications reveal a significantly negative correlation. Given the bias uncovered in Section 3, we suspect that this finding for native displacement is spurious, and instead conclude that no evidence for displacement exists.

2. Native displacement in the existing literature

A straightforward definition of displacement would ask how many native workers (N) respond to the arrival of a single immigrant (F) by leaving their region (state or city) of residence i .⁴ Assuming that the native employment response (ΔN) to an immigrant inflow (ΔF) is linear (at least to a first order approximation), this would imply that the coefficient β in Expression (1) would allow us to identify the presence (if $\beta < 0$) and the magnitude (absolute value of β) of such a phenomenon.

$$\Delta N_{i,t} = \alpha + \beta \cdot \Delta F_{i,t} + u_{i,t} \quad (1)$$

The term $u_{i,t}$ in Eq. (1) captures all the determinants of native employment changes in region i and year t other than the response to immigration inflows. If we allow $\phi(i, t)$ to represent the systematic determinants of native employment changes, the final term would reduce to $u_{i,t} = \phi(i, t) + \varepsilon_{it}$, where ε_{it} is a residual zero-mean random component. By controlling for $\phi(i, t)$, we could directly estimate β from a standard regression of (1). In practice, few if any papers actually employ this direct test of displacement. The regression is likely to be confounded by a number of problems, one being that the average and standard deviation of ΔN and ΔF are likely to be proportional to the total population in the cell, potentially inducing a spurious positive correlation.

⁴ This definition could be further refined to account for arrivals of immigrants who share similar skill characteristics or occupations of natives within regions.

2.1. Card (2001) and Card (2007)

Card (2001) and Card (2007) offer a solution by standardizing native and foreign-born changes by population levels. This allows for well behaved residuals (after also controlling for systematic effects). Card (2001) begins with the identity in (2), which relates the flow (between periods $t - 1$ and t) of native and foreign-born individuals in any observable cell. The variables N and F represent the stock of native and foreign-born workers, while $L = F + N$ is total employment. The superscripts *Est* and *New* refer to established and newly-arrived immigrants, respectively.

$$\left(\frac{L_t - L_{t-1}}{L_{t-1}} \right) = \left(\frac{N_t - N_{t-1}}{L_{t-1}} \right) + \left(\frac{F_t^{Est} - F_{t-1}}{L_{t-1}} \right) + \left(\frac{F_t^{New}}{L_{t-1}} \right) \quad (2)$$

Card (2007) instead adopts a more generalized approach by substituting $F_t = F_t^{Est} + F_t^{New}$ to arrive at the quantitatively equivalent identity in (3).

$$\left(\frac{L_t - L_{t-1}}{L_{t-1}} \right) = \left(\frac{N_t - N_{t-1}}{L_{t-1}} \right) + \left(\frac{F_t - F_{t-1}}{L_{t-1}} \right) \quad (3)$$

The key to the empirical estimation in both papers is that the final term (in either identity) may be causally correlated with the other terms. Card (2001) tests whether newly-arrived immigrants displace natives by regressing $\left(\frac{N_t - N_{t-1}}{L_{t-1}} \right)$ on $\left(\frac{F_t^{New}}{L_{t-1}} \right)$ across a single cross-section of 175 US cities and six occupation groups for the year 1990 (with 1985 representing $t - 1$). Though this precludes him from exploiting the advantages of a panel dataset, his regressions do include dummy variables for cities and occupation groups.⁵ Negative values would imply displacement. Regressions employing various sample selection criteria and instrumental variables techniques, however, find robustly non-negative coefficients ranging from 0.02 to 0.27. Thus, his results argue against a native internal migration response.

The most significant methodological difference between Card (2001) and Card (2007) is that the former tests the effects of newly-arrived immigrants, whereas the latter's interest is in the effects of foreign-born flows in the aggregate. Table 3 of Card (2007) provides estimation results for regressions of Eq. (4) across a single cross-section of US cities (c).⁶

$$\left(\frac{L_{c,t} - L_{c,t-1}}{L_{c,t-1}} \right) = \alpha + \beta_{Card} \cdot \left(\frac{F_{c,t} - F_{c,t-1}}{L_{c,t-1}} \right) + \varepsilon_{c,t} \quad (4)$$

Under this specification, displacement occurs if estimated coefficients are less than one.⁷ Although the migration analysis is not as thorough as in Card (2001), the results again argue against displacement. Estimated coefficients are near two for OLS regressions and near one for IV specifications. None of the estimates is significantly below one.

2.2. Borjas (2006)

Unlike most analyses of internal migration, Borjas (2006) has the distinct quality of being motivated by theory. It begins with a model of labor demand such that native and foreign-born workers are perfect substitutes within skill-region groups. Foreign-born arrivals are assumed to be exogenous and constant over time. Wages respond immediately to the increase in labor supply, whereas native labor supply has a lagged response because it is

⁵ Estimated coefficients appear in the fourth column of his Table 4.

⁶ Since Card (2007) is a single cross-section of cities, but not city-by-occupation cells, dummy variables are not permitted. Regressions do include the log of initial city population as a control.

⁷ An equivalent and perhaps more direct approach would replace the dependent variable with $\left(\frac{N_{c,t} - N_{c,t-1}}{L_{c,t-1}} \right)$ and test whether $\beta < 0$.

Download English Version:

<https://daneshyari.com/en/article/970893>

Download Persian Version:

<https://daneshyari.com/article/970893>

[Daneshyari.com](https://daneshyari.com)