

Freire, Tiago

Conference Paper

City of God Redux: Inequality, Migration, and Violent Crime in Brazil between 1980 and 2000

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Freire, Tiago (2015) : City of God Redux: Inequality, Migration, and Violent Crime in Brazil between 1980 and 2000, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124690>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

City of God Redux: Inequality, Migration, and Violent Crime in Brazil between 1980 and 2000

Tiago Freire

International Business School Suzhou, Xi'an JiaoTong-Liverpool University, No. 111 Ren'ai Road, SIP, Suzhou, Jiangsu, P.R. China, Telephone: +86-0512-81880450; Fax: +86-512-88161899

E-mail: tiago@tiagofreire.com

Running title: Inequality and violent crime in Brazil

Abstract

There is a long-held belief that inequality is a major determinant of violent crime, particularly homicides. Some previous studies suggest that these results hold in the short term only. This could result from measurement error in income inequality.

This study addresses the issue of measurement error in inequality by using the relationship between migration and inequality. Using rainfall shocks and changes in transport costs as exogenous sources of out-migration from rural areas in Brazil between 1980 and 2000, the study shows how migration from rural areas affects income inequality in urban areas. It finds that not only is there a negative and statistically significant relationship between inequality and crime in Brazil, and that the effects are

much larger than previously thought, but also that this relationship holds in the long term.

Keywords: Crime; Inequality; Rural–Urban Migration; Brazil.

JEL Classification Numbers: J61, J15, K42, R10.

I. Introduction

According to the World Bank's World Development Indicators for 2013, Brazil had a Gini index of 55, making it the country with the 14th largest income inequality in the world between 1993 and 2012. According to homicide statistics for 2013 from the United Nations Office on Drugs and Crime (UNODC), Brazil had the 27th largest homicide rate in the world in 2013, with 21.8 homicides per 100,000 people, down from 30.2 in 2002. While the relationship between crime—particularly homicides—and inequality has been studied extensively, the results have been mixed; some authors find evidence of a short-term relationship only. This article examines whether there was a long-term relationship between violent crime (i.e., homicides) and inequality in Brazil between 1980 and 2000.

To understand why a relationship exists between homicides and income inequality, violent crime should be viewed as an activity complementary to other crimes with economic motivation (i.e., property or drug-related crime), as proposed by Grogger (2000). Paulo Lins' award-winning book from 1997, *Cidade de Deus* (City of God), on which a movie by the same name is based, provides a good description of the relationship between violent crime and inequality in Brazilian cities using interviews with residents in a *favela* (shanty town) of Rio de Janeiro. Brazilians living in the *favela* join gangs to sell drugs to make a living. To expand their market and control distribution in an area, gang members often resort to fatal gun-related violence.

While the literature on income inequality and crime records clear evidence of correlation between property crime and inequality, evidence of correlation for violent crime,

particularly homicide, is mixed. For instance, Fajnzylber et al. (2002a, 2002b) found a positive and statistically significant impact of income inequality (using the Gini index) on homicide and robbery rates using a panel of 34–45 countries between 1965 and 1995. Similarly, Scorzafave and Soares (2009) and Sachsida et al. (2010) found a positive and statistically significant correlation between homicide rates and inequality across Brazilian states between 1981 and 1995.

Conversely, Kelly (2000) used data on violent crime (including murder) and property crimes across counties in the US in 1991 and found that while income inequality (measured by the difference between mean and median household incomes) has a positive and statistically significant impact on property crimes, there is no impact on violent crime, in particular murder. Choe (2008) confirmed this result using a panel of US states between 1995 and 2004. Wu and Wu (2012) found a negative and statistically significant relationship between murder and income inequality for the UK.

This study finds a negative and statistically significant relationship between homicides and income inequality in Brazil using data for 1980–2000, a period for which there is consistent data. Consistent with the findings of Scorzafave and Soares (2009) and Sachsida et al. (2010), the results of this study show that an increase of 0.1 in the high to low skill wage ratio leads to a 3.7% increase in the number of homicides in Brazil's urban areas.

Furthermore, with the exception of Choe (2008), few authors have controlled for time fixed effects; thus, they might have identified only short-term and not long-term effects. In fact, Saridakis (2004) found no evidence of a long-run relationship between violent crime (including murder) and income inequality (measured by the Gini index) using time

series data for the US between 1960 and 2000, although a negative short-run relationship exists. Similarly, Brush (2007) employed US census data for 1990–2000, and while he found a positive relationship between inequality and crime in the cross-sectional analysis, the relationship was negative or zero with first differenced data. Similar inconsistent results between inequality and crime were found by Neumayer (2005) for a panel of countries similar to that used by Fajnzylber et al. (2002a, 2002b).

Measurement error in the inequality variable may be one reason for these differences. In particular, Bound and Krueger (1991) and Bound et al. (1994) showed that measurement error could worsen with the inclusion of fixed effects. This problem can be solved using an Instrumental Variable (IV) approach. Fajnzylber et al. (2002a, b) and Scorzaface and Soares (2009) used the dynamic panel data methods proposed by Arellano and Bond (1991) and Blundell and Bond (1998) to achieve identification. These methods use lagged variables as instruments for endogenous variables. However, lagged variables may not solve the measurement error problem if measurement error is correlated over time (Chen et al., 2008; Biørn, 2012; Meijer et al., 2013).

This study shows a long-term relationship between homicide and income inequality. It uses migration as an IV to address the measurement error problem. Using rainfall shocks and changes in transport costs in rural areas as exogenous shocks to rural migration flows into urban areas of Brazil between 1980 and 2000, the problem of measurement error is addressed even if it is correlated over time. The results show that migration is strongly correlated with inequality while being uncorrelated with the error term. Not only is there a positive and statistically significant relationship between inequality and homicides in Brazilian cities between 1980 and 2000, but the result is up to three times

larger than previously thought. This holds in both the short and long run and is robust even with assumptions about the distribution of the probability of criminal behavior, different samples, and an inequality proxy.

Finally, the study shows that there is no direct long-term relationship between migration and homicide in Brazil. To control for all mechanisms through which migration affects homicides, the study follows Bianchi et al. (2012) and Spenkuch (2014) and includes the migration rate directly in regressions. Like these studies, I find that migration has a positive impact on homicides. However, once the IVs on the migration rate into cities and income inequality are used, the results show that migration has only an indirect effect on homicides through changes in inequality, which is consistent with previous studies.

This article is structured as follows. The empirical strategy is developed in Section II. A simple model of criminal behavior is constructed for the relationship between crime and inequality. Next, estimation issues are examined and the relationship between inequality and migration is analyzed. Section III describes the data. Section IV presents the results and develops the IV approach. Section V concludes.

II. Empirical Strategy

The basic economic model of crime is based on the work of Becker (1968) and Ehrlich (1973). An individual will commit a crime if his indirect utility from the criminal activity

is higher than that from working, net of the expected cost of being caught. Therefore, the decision to commit a crime is given by

$$Crime_{irt}^* = 1 \left[\ln Y_{irt}^{crime} - \ln Y_{irt}^{work} + \ln C_{irt} + \eta_{irt} \geq 0 \right] \quad (1)$$

where Y_{irt}^{crime} and Y_{irt}^{work} are the returns to committing crime and legal work, respectively, for individual i in city r at time t , and C_{irt} measures the expected cost of committing a crime, such as the likelihood of being caught and sent to jail. η_{irt} is the error term. $1[\cdot]$ is an indicator function, which takes the value one if the value inside the brackets is positive, and zero otherwise.

To test this theory, a survey is needed of the overall population, including the incomes of individuals who committed crimes (usually incarcerated individuals). However, I am unaware of any survey that meets this requirement for Brazil. Most studies on crime rely on the area approach, in which estimates are obtained by comparing crime rates across regions (and potentially across time). Then, the crime rate in a particular city r for a particular year t is $E \left[Crime_{irt}^* \right] = (Crime\ rate)_{rt}$.

As noted by Stoker (2008) and Durlauf et al. (2010), to obtain the expected value of crime across regions, the distribution of the probability to commit a crime needs to be assumed. A common assumption for the probability of criminal behavior is a Poisson distribution (Kelly, 2000; Osgood, 2000). Accordingly, Equation 2 can be estimated by the Generalized Method of Moments (GMM):

$$\ln(Crime_{rt}) = \alpha + \gamma [\ln \bar{Y}_{rt}^{crime} - \ln \bar{Y}_{rt}^{work}] + \beta \ln \bar{C}_{rt} + \varepsilon_{rt} \quad (2)$$

Thus, for city r at time t , $(Crime\ rate)_{rt}$ is the number of crimes per capita, and $\ln \bar{Y}_{rt}^{work}$ is the average log income from work; $\ln \bar{Y}_{rt}^{crime}$ is the average log income from crime; while $\ln \bar{C}_{rt}$ is the average of the log of the cost of committing a crime. α is a constant.

Since income from labor and income from crime are not observed simultaneously, individual and city characteristics X_{irt} are used to determine income from both sources, such as education and experience. Another determinant of potential earnings from criminal behavior is the other individual's work income $Y_{-i,rt}^{work}$ as a measure of the maximum potential gain from criminal activity, $\ln \bar{Y}_{rt}^{crime}$, in city r at time t . Ehrlich (1973) and Kelly (2000) use income inequality $Ineq_{rt}$ as a proxy for potential gains from criminal behavior. Therefore, $Y_{-i,rt}^{work} = Ineq_{rt} - u_{rt}$, where u_{rt} is a general term for measurement error, which has an expected value of zero and is uncorrelated with other observables and the error term in the equation. Most studies estimate a form of Equation 3.

$$\ln(Crime_{rt}) = \gamma_1 \bar{X}_{rt} + \gamma_2 Ineq_{rt} + \beta \ln \bar{C}_{rt} + (\varepsilon_{rt} - \gamma_2 u_{rt}) \quad (3)$$

where \bar{X}_{rt} is the average of characteristics of individuals in city r at time t , as well as characteristics of the city, and $Ineq_{rt} = Y_{-i,rt}^{work} + u_{rt}$ is a measure of income inequality. As noted by Stoker (2008) and Durlauf et al. (2010), the response of crime rates to individual level characteristics X_{irt} using aggregate individual data cannot be recovered, although the response of crime to city characteristics, such as the impact of income inequality on crime, can be estimated.

To see how the use of a proxy variable affects estimates of γ_2 , this study ignores the other variables in Equation 3.

$$\ln(Crime_{rt}) = \gamma_2 Y_{-i,rt}^{work} + \varepsilon_{rt} \quad (4)$$

Then, the estimate of γ_2 will be

$$\begin{aligned} \hat{\gamma}_2 &= \frac{\text{cov}(Ineq_{rt}; \ln Crime_{rt})}{\text{var}(Y_{-i,rt}^{work})} \\ &= \frac{\text{cov}(Y_{-i,rt}^{work} + u_{rt}; \gamma_2 Y_{-i,rt}^{work} + \varepsilon_{rt})}{\text{var}(Y_{-i,rt}^{work} + u_{rt})} \\ &= \gamma_2 \frac{\sigma_{Y_{-i,rt}^{work}}}{\sigma_{Y_{-i,rt}^{work}} + \sigma_{u_{rt}}} \\ &= \gamma_2 \frac{1}{1 + \frac{\sigma_{u_{rt}}}{\sigma_{Y_{-i,rt}^{work}}}} \end{aligned} \quad (5)$$

where σ_m is the standard error of m . The estimates of γ_2 will be biased towards zero.

Furthermore, as pointed out by Griliches and Hausman (1986) and Pischke (2007), standard errors will be biased upward, leading to lower t -statistics.

The inclusion of fixed effects, such as city fixed effects, can worsen measurement errors.

Like Bound and Krueger (1991), Bound et al. (1994), and Pischke (2007), this study considers the case of city fixed effects α_r only.

$$\ln(Crime_{rt}) = \alpha_r + \gamma_2 Y_{-i,rt}^{work} + (\varepsilon_{rt} - \gamma_2 u_{rt}) \quad (6)$$

Taking the first difference of the data gives a similar regression model to before.

$$\Delta \ln(Crime_{rt}) = \gamma_2 \Delta Y_{-i,rt}^{work} + (\Delta \varepsilon_{rt} - \gamma_2 \Delta u_{rt}) \quad (7)$$

The estimate of γ_2 will be biased as follows.

$$\begin{aligned}
\hat{\gamma}_2 &= \gamma_2 \frac{\sigma_{\Delta Y_{-i,rt}^{work}}}{\sigma_{\Delta Y_{-i,rt}^{work}} + \sigma_{\Delta u_{rt}}} \\
&= \gamma_2 \frac{\sigma_{Y_{-i,rt}^{work}}(1-\rho)}{\sigma_{Y_{-i,rt}^{work}}(1-\rho) + \sigma_{u_{rt}}(1-r)} \\
&= \gamma_2 \frac{1}{1 + \frac{\sigma_{u_{rt}}(1-r)}{\sigma_{\Delta Y_{-i,rt}^{work}}(1-\rho)}}
\end{aligned} \tag{8}$$

where ρ is the autocorrelation of the maximum potential gain from criminal activity

$Y_{-i,rt}^{work}$ over time, and r is the autocorrelation of the measurement error term u_{rt} . It is

easier to see that when the measurement error is not correlated over time, $\rho \approx 0$, and

when $Y_{-i,rt}^{work}$ is highly correlated over time, $r \approx 1$; adding fixed effects increases the bias.

A possible solution is to find an instrument Z_{rt} that is uncorrelated with the measurement

error $cov(\Delta u_{rt}; Z_{rt})=0$ but is correlated with changes in the maximum potential gain from

criminal activity $cov(\Delta Y_{-i,rt}^{work}; Z_{rt}) \neq 0$. Thus, the estimate of γ_2 is given by

$$\begin{aligned}
\hat{\gamma}_2 &= \frac{cov(\Delta \ln Crime_{rt}; Z_{rt})}{cov(\Delta Ineq_{rt}; Z_{rt})} \\
&= \frac{cov(\gamma_2 \Delta Y_{-i,rt}^{work} + \varepsilon_{rt}; Z_{rt})}{cov(\Delta Y_{-i,rt}^{work} + u_{rt}; Z_{rt})} \\
&= \gamma_2 \frac{cov(\gamma_2 \Delta Y_{-i,rt}^{work}; Z_{rt})}{cov(\Delta Y_{-i,rt}^{work}; Z_{rt})} \\
&= \gamma_2
\end{aligned} \tag{9}$$

Therefore, I can obtain unbiased estimates of γ_2 . The standard instrument used in the literature is the second lag of the level of income inequality, as suggested by Bound et al. (1994), Arellano and Bover (1995), and Bond (2002). However, this method will provide unbiased estimates of γ_2 only if the error term $\Delta \varepsilon_{rt} - \gamma_2 \Delta u_{rt}$ is not serially correlated. As Durlauf et al. (2010) note, there is no reason to assume that this is true, since individuals may be forward looking and the nature of individual- and city-specific heterogeneity is unknown. In fact, Chen et al. (2008) and Biørn (2012) use Montecarlo simulations to show that under a general structure for the measurement error term u_{rt} Bound et al.'s (1994) and Arellano and Bover's (1995) method will still provide biased estimates of γ_2 .

In this study, the proxy for income inequality is the ratio of high skill (H) (i.e., with 12 years of education or more) median hourly wages $\frac{w_{Hrt}}{w_{Lrt}}$ in city r in year t to those of low skill (L) skill (i.e., with less than 12 years of education). This depends on the supply of high and low skill workers $\frac{N_{Hrt}}{N_{Lrt}}$, as noted by Borjas (2006), Card (2009), Freire (2010), and Peri (2011).

The 1991 and 2000 Brazilian population censuses did not track the number of migrants by municipality but recorded only recent cases of migration (people who moved to a city in the five years prior to the census year). Thus, only migrant flow can be measured and not the stock of migrants. When individuals from rural areas move to city r , the relative

supply of high to low skill workers in that city $\Delta \left[\frac{N_{Hrt}}{N_{Lrt}} \right]$ changes, which in turn changes

the ratio of high to low skill workers' wages $\Delta \left[\frac{w_{Hrt}}{w_{Lrt}} \right] = \Delta Ineq_{rt}$

It should be pointed out that migration flows are themselves endogenous, with changes in urban wages promoting migration from rural areas. As explained in Section IV, this study finds that rainfall shocks in rural areas and changes in transport costs are important exogenous determinants of out-migration from rural areas for low skill migrants and of out-migration from rural areas to urban areas, respectively, and can be used to construct an exogenous IV.

III. Crime, Income Inequality, and Rural–Urban Migration in Brazil

The three sources of crime statistics for Brazil include data on homicides from the public health database, police records for some states, and data from victimization surveys (Santos and Kassouf, 2008). Data on victims are very limited; there is only one nationally representative survey, and the existing panel data are limited to four states (World Bank, 2006). Police records prior to 2002 do not exist at the national level (World Bank, 2006). According to the 2002 police records, the public health database (DATASUS) over reports the number of homicides (World Bank, 2006); while public health records count homicide deaths resulting from legal interventions (killings by police and public security forces), war, and declared homicides. Therefore, this study's

estimates are upper bounds of the effect of income inequality on homicides. Regardless of the data source, Brazil has one of the highest homicide rates in the world, according to UNODC (2009), and it has been increasing over time (World Bank, 2006).

Given these data limitations, following Scorzafave and Soares (2009) and Sachside et al. (2010), this study uses homicides reported in DATASUS as a proxy for criminal activity. As explained in Section II, this study assumes that homicides are committed in relation to property and drug-related crimes, and therefore, they can be studied with an economic model of crime, which is consistent with the findings of Fajnzylber et al. (2002a, 2002b). This study assumes that the homicide rate is a measure of the number of crimes committed, and therefore, follows a Poisson distribution. The regressions measure the response to changes in income inequality at the intensive and extensive margins of criminal behavior. However, as noted in Section IV, the conclusions hold if a uniform distribution is assumed and the crime rate measures the number of criminals in a city, the extensive rate.

The data are restricted to the census years 1980, 1991, and 2000, since a definition of “city” is required that allows for a comparison of city characteristics over time. This study uses the 123 urban agglomerations (also referred to as “cities” or “urban areas”) defined by Mata et al. (2007) as the metropolitan statistical areas of the US, allowing city-level comparisons between 1980 and 2000. There is no such definition of “city” that allows for the inclusion of the 2010 census data. The basic statistics appear in Table 1. The results are consistent with the current literature: the number of homicides in these 123 cities increased 73% during this time period (23.5 per 100,000 people in 1980 to 40.59 per 100,000 in 2000).

According to the National Penitentiary Department Database of the Ministry of Justice (Infopen, 2008), most incarcerated people are males under the age of 45 years (Table 2), and most incarcerated males have had a high school education or less (Table 3). These characteristics are also assumed to be the characteristics of criminals committing homicides, assuming that the Brazilian justice system is as efficient in catching and convicting murderers as individuals committing other crimes. Therefore, Table 1 also reports the number of homicides per 100,000 committed by young men (between 15 and 45 years) with low skill (less than 12 years of education).

According to Cerqueira (2010), between 1980 and 1991, more police officers were employed in response to the increasing number of homicides. Using data of the 1980, 1991, and 2000 Brazilian population censuses, the number of people working in public security (police and military police) is calculated. The number of police officers in urban areas reported in the censuses increased 31% from 375.78 per 100,000 in 1980 to 492.91 per 100,000 in 1991 (Table 1). Cerqueira (2010) notes that this did not curb crime due to increasing inefficiency in investigating and convicting criminals. Between 1990 and 2000, although government spending on public safety increased, Cerqueira (2010) notes a 41.7% decrease in the size of the police force in urban areas (Table 1). This period also saw a significant increase in the size of the private security industry, usually staffed with workers from the public security sector (which creates an obvious incentive incompatibility problem).

The period 1980 to 2000 saw other changes in the Brazilian economy, which may have affected criminal behavior patterns. In response to the country's balance of payments

crisis, wages decreased between 1980 and 1991 for both high skill (high school graduates or higher) and low skill (high school dropouts or lower) workers (Table 1).

Driven in part by government budget deficits, unemployment increased slightly between 1980 and 1991.¹ However, due to the reforms of the Real Plan, the number of unemployed men aged 15–45 years increased significantly. In particular, Table 1 shows that the percentage of young and low skill men outside the labor force increased during this period.

Income inequality increased between 1980 and 1991, with hourly wages dropping more for low skill workers than high skill workers. Between 1991 and 2000, hourly wages rebounded, with wages for low skill workers increasing more than those for high skill workers, leading to a decrease in income inequality (Table 1). The relationship between homicide rates and the ratio of high to low skill wages across cities appears in Fig. 1. As predicted by my model, there is a positive relationship between changes in income inequality and changes in the homicide rate. However, as before, there is large variation across cities, consistent with the possibility of measurement error.

As explained in Section II, changes in the relative supply of high and low skill workers in cities, driven by rural migration, affects urban income inequality. Therefore, it is not

¹ While the 1980 and 1991 Brazilian population censuses queried if individuals were employed, looking for a job, or outside the labour force, this information was not collected in the 2000 census. Therefore, this study measures unemployment as those individuals without jobs ($1 - \text{the labour force participation rate}$). This value could be affected by the number of young people who decide to continue studying but has remained relatively stable over the given period.

surprising that there is a positive relationship between migration and income inequality (Fig. 2) and between migration and the homicide rate (Fig. 3).

IV. Results

Equation 3 in first difference is estimated using GMM with clustered-robust standard errors.

$$\Delta \ln(\text{Crime}_{rt}) = \alpha_t + \gamma_1 \Delta \bar{X}_{rt} + \gamma_2 \Delta \text{Ineq}_{rt} + \beta \Delta \ln \bar{C}_{rt} + e_{rt} \quad (10)$$

where the coefficient of interest γ_2 measures the impact of income inequality

$\text{Ineq}_{rt} = \frac{w_{Hrt}}{w_{Lrt}}$ on the crime rate. \bar{X}_{rt} is a set of city characteristics that is standard in the

literature. These characteristics include unemployment rate and hourly wage. In some specifications, variables that are standard in migration- and crime-related literature are added, namely, city size and the fraction of recent migrants as a percentage of lagged city size. To measure city size, this study uses the lagged size of the resident population, which includes only individuals who were living in the city in the five years prior to the previous Census, and therefore, it is predetermined and uncorrelated with current or lagged city characteristics in the error term. α_t denote a set of year dummies. Two versions of the model are estimated, one without year dummies (similar to the specification standard in the literature) and the other with year dummies (to capture the long-run effects of income inequality on the crime rate). $\ln \bar{C}_{rt}$ is proxied by the size of the police force in city r in year t .

As discussed in Section II, we also know that individual characteristics such as gender, education, and age are important determinants of the crime rate. However, including a control for these individual characteristics is likely to bias my estimates (Stoker, 2008 and Durlauf et al., 2010). Therefore, I restrict the measures of the crime rate, unemployment, and income to a sample of young men. In Section IV, I show that my results remain unchanged when I extend my sample to all men and women between the ages of 15 and 65.

Basic regression

The model in Equation 10 is estimated by including only income inequality (changes in the ratio of high to low skill hourly wages), wages (changes in the average log hourly wage for young and low skill men), unemployment (changes in the young and low skill male unemployment rate), and growth of the police force; the specification is similar to that used by Scorzafave and Soares (2009) and Sachsida et al. (2010) for Brazil.

The results appear in columns 1 and 2 of Table 4. Like the work of Sachsida et al. (2010) for Brazil, the model shown in column 1 does not control for changes across time (no constant). In this specification, which is consistent with Sachsida et al. (2010), inequality has a positive and statistically significant impact on the growth rate of homicides. In particular, an increase of 0.1 in the ratio of high to low skill wages leads to a 1.3% increase in the number of homicides. Also consistent with Sachsida et al. (2010), an increase in unemployment leads to a positive and statistically significant increase in the number of homicides, an increase in wages leads to a decrease in the number of homicides, while the number of police officers has no impact on crime. Column 2

includes time dummies. These results are comparable with studies about the long-term impact of inequality on crime, such as Saridakis (2004) and Neumayer (2005). While inequality has a positive impact on the number of homicides, it is not statistically significant; this is consistent with the idea that the impact of inequality on crime is only a short-term relationship. Furthermore, the point estimates for changes in the high to low skill wage ratio are smaller than those presented in column 1, which is consistent with the possibility of measurement error being worse when year fixed effects are included.

To use migration as an instrument, the study must control for all mechanisms through which migration may affect the homicide rate. The issue is that migration does not play a direct role in the basic microeconomic model of criminal behavior. However, migration might affect crime through indirect channels. For instance, in line with the migration literature (LaLonde and Topel, 1991; Ottaviano and Peri, 2012), Bianchi et al. (2012) argue that migrants might face different earnings potential from legal work and crime. Borjas et al. (2010) argue that migration affects crime rates through the labor market conditions of local residents, namely wages and unemployment, which is also consistent with the migration literature (Borjas, 2003). Finally, Bianchi et al. (2012) and Bell et al. (2013) observe the possibility that migrants have a different propensity to commit crimes than local residents, and therefore, migration captures changes in the population composition. Therefore, this study adds migration directly as part of $\Delta\bar{X}_{it}$ in Equation 10 as a proxy for all mechanisms through which migration may affect crime, including income, unemployment, and changes in the population composition. It also includes a control for city size, which is standard in the migration literature. This is shown in columns 3 and 4 of Table 4. The results are consistent with those of Bianchi et al. (2012).

In particular, regardless of whether controls for changes over time are included or not, inequality has a positive but statistically significant impact on the number of homicides. Furthermore, other individual characteristics, such as unemployment and wages, as well as city characteristics, such as city size growth and growth of the police force, are not statistically significant in either model. Consistent with Bianchi et al. (2012), only the coefficient for the fraction of recent young and low skill migrants is positive and statistically significant in both models.

Finally, in columns 5 and 6 of Table 4, this study estimates a reduced model with only inequality and fraction of migrants. The impact of inequality (and migration) is not statistically different from that reported in columns 3 and 4; notably, inequality (measured by changes in high to low skill wages) has a positive but not statistically significant impact on the number of homicides.

As argued in Section II, any measure of income inequality is likely to have measurement errors, which can be addressed with an IV approach.

Instrumental variables

This study constructs an IV using a strategy similar to a supply–push IV (SPIV), developed from Freire (2010). The decision to leave rural areas is separated from the decision of where to migrate. To explain migration from rural areas, this study uses rainfall shocks and improvements in the transport network in rural areas of Brazil between 1980 and 2000, which affect only rural areas. Distance between the origin (rural municipality) and destination (city) is used to explain the decision of where to settle. The

study shows that our instruments allow for variation in migration patterns across groups (high and low skill, and men and women) to build two instruments for income inequality and the proportion of recent migrants from the predicted values of the migration models.

How many people decide to leave rural areas. Table 5 shows the characteristics of rural areas. The average number of people leaving a rural area decreased by 2.3% between 1991 and 2000 (not shown in the table), but the composition changed, with more high skill workers leaving in 2000 than in 1991. We observe the opposite pattern for low skill workers.

From the distribution of workers across sectors (Table 5), we see that in 1991, 43.62% of low skill men worked in farming, dropping to 30.45% in 2000. Therefore, it is likely that individuals in rural areas working in agriculture (or businesses complementary to agriculture) would respond to rainfall shocks (droughts or floods), which are likely to affect income from agricultural production.

The number of high skill people increased by 43% between 1991 and 2000, much faster than the number of low skill people living in rural areas (12%; Table 5), but the percentage of high skill people migrating remained the same at 48%. Furthermore, we notice that the average cost of moving, as measured by the index of transport cost from rural municipalities to São Paulo, as constructed by Castro (2002), dropped by 14% between 1991 and 2000.² Therefore, as noted by Vidal (1998), Docquier and Rapoport (2004), and Beine et al. (2008), the possibility of migration (due to reduced transport

² This index is the result of a linear programming exercise using information on the conditions of the roads connecting the rural municipality and São Paulo.

costs) leads to more investment in education, and thus, increased out-migration of high skill people from rural areas. Furthermore, it is unlikely that more high skilled people are leaving rural areas in search of higher levels of education as the number of people with a college degree remained largely unchanged during this period (Souza, 2001).

Our basic regression becomes

$$\ln Migrants_{i,rural,t} = \delta_0 + \delta_1 \ln N_{i,rural,(t-10)} + \delta_2 Rain_{rural,t} + \delta_3 \ln Transp_{rural,t-5} + \delta_4 W_i + v_{i,rural,t} \quad (11)$$

where $N_{i,rural,(t-10)}$ is the (lagged) number of people living in a rural area, *rural*, in the previous census year ($t-10$) with high or low skill i ; $Rain_{rural,t}$ is the (log) average of monthly rainfall (mm) in *rural* in year t (since the timing of the drought and its impact is uncertain, lagged rainfall is also included)³; and $\ln Transp_{rural,t-5}$ is the (log) index of the transport cost to São Paulo from *rural* in 1986–1995, ($t-5$). W_i is a set of controls depicting the characteristics of the rural area of origin (including log agricultural area, year dummies, and municipality fixed effects).

Since there are a considerable number of rural municipalities with zero out-migration (Table 6; around 70% for high skill migration) and fixed effects are included, this equation is estimated using non-parametric methods developed by Honore (1992).

Table 6 presents the results for high and low skill men by group. Since rural municipalities' fixed effects are included, our coefficients are interpreted as responses to shocks (deviations from the average over the period). As in Freire (2010), the study finds

³ A quadratic term was initially included in our specifications, but was later dropped, as it was statistically insignificant.

that reductions in transport costs cause increased out-migration of high skill people; for men only, a 10% decrease in transport costs increases the number of high skill migrants by 3%. Furthermore, rainfall shocks affect migration of low skill people only, in particular, a 1 SD decrease in rainfall leads to an increase of 5% in migration. These results do not hold for all of Brazil. In the drought-prone area in the northeast, the impact of drought differs. As noted by Baer (2008), this region receives government aid, which is often misused, in years of drought. Therefore, during years of plenty rainfall (when there are no government transfers), out-migration decreases for both low and high skill people. Finally, the instruments seem to be relevant for explaining out-migration, passing the non-linear version of an F -test on all coefficients equal to zero.

Where migrants go. Distance between the destination and origin is used to explain the historical migrant settlement pattern. The 1991 Brazil population census asks the question: “Where were you living 10 years ago?” Along with information on “When did you move to your current municipality,” this study uses the migrant settlement pattern between 1981 and 1985 as an historical pattern of migrant settlement across cities. Furthermore, distance between the origin and destination is used as a proxy for moving cost; therefore, it is orthogonal to changes in conditions in urban areas between 1985 and 1999.

This study constructs a cross section of migrants in 1981–1985 from 3,207 rural origins to 123 urban destinations. Pooling high and low skill men, the following regression is run.

$$\frac{Migrants_{i,rural,urban}}{\sum_c Migrants_{i,rural,c}} = \eta_0 + \eta_1 Distance_{rural,urban} + \eta_2 W_{i,rural,urban} + \varpi_{i,rural,urban} \quad (12)$$

where $\frac{Migrants_{i,rural,urban}}{\sum_c Migrants_{i,rural,c}}$ is the share of migrants moving from a rural area to an urban area c for each group (high or low skill) i . The measure of $Distance_{rural,urban}$ is the great-circle distance between the geographical centers of the rural municipality of origin and the destination city. $W_{i,rural,urban}$ are controls for different group characteristics.

Our results, like Freire (2010), show that migrants are more likely to move to cities closer to their home (rural municipality; column 1 of Table 7). In particular, a city that is 10% closer receives 0.2% more migrants. Column 2 of Table 7 shows that distance matters less for high skill men. In columns 3 and 4, I check whether this result can be attributed to the fact that municipalities nearer cities have a higher supply of high or low skill men, by controlling for how many people live in the rural and urban areas of origin. Even with this control, distance remains statistically significant.

Next, the two estimates (how many people migrate from rural areas and where they decide to go) are combined to construct the instruments.

Results using instrumental variables

As in Freire (2010), this study builds an exogenous migration shock for each city (using estimates from the previous regression), as follows.

$$\left(\Delta\hat{N}_{i,c,t}\right)^{mig} = \sum_{j=t-5}^t \sum_{rural} \left(\frac{\hat{Migrants}_{i,rural,urban}}{\sum_c \hat{Migrants}_{i,rural,c}} \right) \hat{Migrants}_{i,rural,j} \quad (13)$$

where $\hat{Migrants}_{i,rural,j}$ is the predicted number of male migrants of high and low skill i , leaving rural area *rural* at year t , using the results from Table 6. $\left(\frac{\hat{Migrants}_{i,rural,urban}}{\sum_c \hat{Migrants}_{i,rural,c}} \right)$ is the predicted fraction of total male migrants of high and low skill i from *rural*, who will settle in urban area c , using results from Table 7. Therefore, $\left(\Delta\hat{N}_{i,c,t}\right)^{mig}$ will be an exogenous shock to the high and low skill i population living in city c in year t , which depends only on characteristics of individuals and municipalities of origin (rural areas) and is, therefore, orthogonal to the characteristics of cities and individuals living in them.

Therefore, $\ln\left(\Delta\hat{N}_{i,c,t}\right)^{mig}$ and $\frac{\left(\Delta\hat{N}_{H,c,t}\right)^{mig}}{\left(\Delta\hat{N}_{L,c,t}\right)^{mig}}$ are used as instruments for the change in high to low skill wage ratio $\Delta Ineq_{rt}$ and the fraction of recent migrants in Equation 10.

The first-stage results for inequality appear in columns 1 (no constant) and 2 (time dummies) of Table 8. The results show that changes in the ratio of high to low skill migrants are correlated with the endogenous variables and changes in high to low skill wages (high t -statistics and high F -statistics). Moreover, the instruments predict the log of recent migrants. In column 1, an increase in the predicted values of high to low skill young and low skill men decreases high to low skill wages. However, for column 2, the

sign of the coefficient for the predicted changes in the high to skill ratio is positive, contrary to expectations, although the sign for the percentage of recent young and low skill male migrants is negative. This is because cities with a higher ratio of high to low skill migrants also received more migrants (Table 7).

The first-stage results for migration are reported in columns 3 (no constant) and 4 (time dummies) of Table 8. In column 3, the coefficient for the predicted log number of recent migrants is positive and statistically significant (the F -statistic for both instruments is well above 10). However, when time dummies are included in column 4, only changes in high to low skill young and low skill men have a statistically significant impact. In particular, a decrease in the ratio of high to low skill migrants increases the percentage of recent young and low skill male migrants. However, the F -test is less than 10, suggesting the presence of weak instruments. However, the estimated standard errors are only efficient under homoskedasticity. Aggregating individual data leads to heteroskedasticity. Therefore, the second-stage regressions calculate the Stock–Wright Lagrange Multiplier (LM) S -statistic for weak instruments, which is robust to heteroskedasticity (Baum et al., 2007).

The second-stage results for the reduced model appear in columns 1 and 2 of Table 9. When a control for changes common to all cities over time (a model with no constant in column 1) is excluded, inequality has a negative but not statistically significant impact on the number of homicides. Migration has a positive and statistically significant impact on the number of homicides. This is likely caused by weak instruments, as the Stock–Wright test is rejected only at the 10% level. However, after including time dummies, the Stock–Wright test is rejected at the 5% level, and the Kleibergen–Paap test for under-

identification (weak correlation between the instruments and endogenous variables; see Baum et al., 2007) is rejected at the 1% level. In this case, changes in the high to low skill wage ratio (inequality) have a positive and statistically significant impact on the number of homicides. In particular, an increase of 0.1 in the high to low skill wage ratio leads to a 3.7% increase in the number of homicides. This is eight times larger than the coefficient found in column 6 of Table 4 and is consistent with the presence of measurement error. This is consistent with the findings of Bianchi et al. (2012) that migration does not have a statistically significant impact on the number of homicides.

In columns 3 and 4 of Table 9, the model with all other control variables is estimated. Both specifications pass the tests for weak instruments and under-identification. In both columns, changes in the high to low skill wage ratio (inequality) lead to positive and statistically significant increases in the number of homicides. In particular, an increase of 0.1 in the high to low skill wage ratio leads to a 4.25–5.79% increase in the number of homicides. Migration does not have a statistically significant impact on the number of homicides.

In summary, when both instruments are valid, there is a positive and statistically significant relationship between inequality and homicides.

Robustness checks

In columns 5 and 6 of Table 9, the same regressions with all individuals (men and women) between the ages of 15 and 65 years are run. Therefore, this study's measures of inequality, migration, poverty, and unemployment relate to this sample. The results are

consistent with previous results. In particular, a change of 0.1 in the high to low skill wage ratio leads to a 3.6–4.3% (statistically significant) increase in the number of homicides. This result is true only when the instruments are appropriate (i.e., they pass the Stock–Wright test for weak instruments). Migration is never statistically significant. In addition, the results are robust to different measures of inequality. The ratio of the 75th to 25th wage quantile is used as the measure of inequality in columns 7 and 8 of Table 9. As before, when the instruments are valid (the Stock–Wright *S*-statistic is significant at the 5% level), there is a positive and statistically significant relationship between inequality and homicides, while migration does not have a statistically significant impact on homicides.

Finally, the results are robust to the assumption about the distribution of the probability to commit a crime. The online appendix presents the results for a uniform distribution.

V. Conclusion

I began by replicating the findings of previous researchers, who have shown a positive and statistically significant relationship between inequality and homicides in Brazil in the short run. However, many previous studies found no impact in the long run. I explained how these results could be biased downward due to measurement error, a problem that worsens with the inclusion of time dummies. This study showed that this problem cannot be addressed using the standard dynamic panel data models used in the literature. Instead, it used the relationship between inequality and migration to construct an instrument for inequality.

This study showed how rainfall shocks, particularly droughts, in rural areas and changes in transport costs are important determinants of high and low skill migration from rural to urban areas. Furthermore, I argued that the migration flows predicted from these models are orthogonal to the current characteristics of urban areas; therefore, they can be used as instruments. I showed how the instruments are correlated with inequality and the percentage of recent migrants.

The results for instrumental variables show that inequality has a positive and statistically significant impact on homicides in the long run. In particular, it increases the number of homicides. These results are robust to the model specification of this study, the sample, the proxy for inequality, and the assumption made about the probability distribution of committing a crime. Furthermore, the study confirmed the relationship between migration and crime. Notably, it found no statistically significant relationship between migration and the homicide rate, which is consistent with Bianchi et al. (2012) and Spenkuch (2014).

The award-winning book *Cidade de Deus*, which is based on real events and was ultimately adapted as a movie, gives us an idea of how the mechanism between inequality and homicides may work. In this book, migrants from rural areas who settled in the *favela* deal in drugs as a way to escape poverty. They sell drugs to middle-income people living in the city's center. Rivalries between drug dealers in the *favelas* often result in homicides.

This study's results have important implications for public policy on social programs to reduce criminality in Brazil. In particular, since the impact of inequality on homicides is significantly larger than previously thought, social programs with the potential to reduce

inequality, such as *Bolsa Familia*, which started in 2003, are likely to be effective. Furthermore, policies that decrease rural–urban migration may help reduce violent crime in Brazilian cities.

References

- Arellano, Manuel, and Stephen Bond. 1991. “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.” *The Review of Economic Studies* 58, no. 2: 277–97.
- Arellano, Manuel, and Olympia Bover. 1995 “Another Look at the Instrumental Variable Estimation of Error Components Models.” *Journal of Econometrics* 68, no. 1: 29–51.
- Baer, Werner. 2008. *The Brazilian Economy: Growth and Development*, 6th edition. Westport, CT: Lynne Rienner Publishers.
- Baum, Christopher F.; Mark E. Schaffer; and Steven Stillman. 2007. “Enhanced Routines for Instrumental Variables/GMM Estimation and Testing.” *The Stata Journal* 7, no. 4: 465–506.
- Becker, Gary S. 1968. “Crime and Punishment: An Economic Approach.” *The Journal of Political Economy* 76, no. 2: 169–217.
- Beine, Michel; Frederic Docquier; and Hilel Rapoport. 2008. “Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers.” *The Economic Journal* 118, no. 528: 631–652.

- Bell, Brian; Francesco Fasani; and Stephen Machin. 2013. "Crime and Immigration: Evidence from Large Immigrant Waves." *Review of Economics and Statistics* 21, no. 3: 1278–1290.
- Bianchi, Milo; Paulo Buonanno; and Paolo Pinotti. 2012. "Do Immigrants Cause Crime?" *Journal of the European Economic Association* 10, no. 6: 1318–1347.
- Biørn, Erik. 2012. The Measurement Error Problem in Dynamic Panel Data Analysis: Modeling and GMM Estimation. Memorandum no. 02/2012. Department of Economics, University of Oslo, Norway.
- Blundell, Richard, and Steven Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87, no. 1: 115–143.
- Bond, Richard. 2002. "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice." *Portuguese Economic Journal* 1, no. 2: 141–162.
- Borjas, George. 2003. "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics*, 118, no. 4: 1335–1374.
- Borjas, George. 2006. "Native Internal Migration and the Labor Market Impact of Immigration." *The Journal of Human Resources* 41, no. 2: 221–58.
- Borjas, George J.; Jeffrey Grogger; and Gordon H. Hanson. 2010. "Immigration and the Economic Status of African-American men." *Economica* 77, no. 306: 255–282.
- Bound, John, and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9, no. 1: 1–24.

- Bound, John; Charles Brown; Greg Duncan; and Wilard Rodgers. 1994. "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *Journal of Labour Economics* 12, no. 3: 345–368.
- Brush, Jesse. 2007. "Does Income Inequality Lead to More Crime? A Comparison of Cross-Sectional and Time-Series Analyses of United States Counties." *Economic Letters* 96, no. 2: 264–68.
- Card, David. 2009. "How Immigration Affects US Cities." In *Making Cities Work: Prospects and Policies for Urban America*, ed. Robert P. Inman. Princeton: Princeton University Press, 158–200.
- Castro, Newton. 2002. "Custos De Transporte E Produção Agrícola No Brasil, 1970–1996." *Agricultura em São Paulo* 49, no. 2: 87–109.
- Cerqueira, Daniel R.C. (2010) "*Causas e Consequencias do Crime do Brasil*," PhD diss., Pontifica Universidade Catolica do Rio de Janeiro.
- Chen, Jie; Shawn Ni; and Michael Podgursky. 2008. "Estimating Dynamic Panel Data Models with Measurement Errors with an Application to School Evaluation based on Student Test Scores," Proceedings of the 2008 meeting of the American Statistical Association, pp. 951–957.
- Choe, Jongmook. 2008. "Income Inequality and Crime in the United States." *Economics Letters* 101, no. 1: 31–3.
- Docquier, Frederic, and Hilel Rapoport. 2004. *Skilled Migration: The Perspective of Developing Countries*. World Bank Policy Research Working Paper no. 3382, New York.

- Durlauf, Steven; Salvador Navarro; and David Rivers. 2010. "Understanding Aggregate Crime Regressions." *Journal of Econometrics* 158, no. 2: 306–317.
- Ehrlich, Isaac. 1973. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy* 81, no. 3: 521–565.
- Fajnzylber, Pablo; Daniel Lederman; and Norman Loayza. 2002a. "Inequality and Violent Crime." *The Journal of Law and Economics* 45, no. 1: 1–40.
- Fajnzylber, Pablo; Daniel Lederman; and Norman Loayza. 2002b. "What Causes Violent Crime?" *European Economic Review* 46, no. 7: 1323–1357.
- Freire, Tiago, (2010) "Essays in Applied Economics." PhD diss., Brown University, Providence.
- Griliches, Zvi, and Jerry A. Hausman. 1986. "Errors in Variables in Panel Data." *Journal of Econometrics* 31, no. 1: 93–118.
- Grogger, Jeff. 2000. "An Economic Model of Recent Trends in Violence." In *The Crime Drop in America*. Cambridge, UK: Cambridge University Press, pp. 266–287.
- Honore, Bo E. 1992. "Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects." *Econometrica* 60, no. 3: 533–565.
- Infopen. 2008. *Infopen – Informacoes Penitenciarias*. www.infopen.gov.br (accessed November 5, 2013).
- Kelly, Morgan. 2000. "Inequality and Crime." *The Review of Economics and Statistics* 82, no. 4: 530–539.

- LaLonde, Robert J., and Robert H. Topel. 1991. "Immigrants in the American Labor Market: Quality, Assimilation, and Distributional Effects." *American Economic Review* 81, no. 2: 297–302.
- Mata, Daniel; Uwe Deichmann; John Vernon Henderson; Somik Lall; and Hyoung Wang. 2007. "The Determinants of City Growth in Brazil." *Journal of Urban Economics* 62, no. 2: 252–272.
- Meijer, Erik; Laura Spierdijk; and Tom Wansbeek. 2013. "Measurement Error in the Linear Dynamic Panel Data Model." In *ISS–2012 Proceedings Volume On Longitudinal Data Analysis Subject to Measurement Errors, Missing Values, and/or Outliers*. New York: Springer, pp. 77–92.
- Neumayer, Eric. 2005. "Inequality and Violent Crime: Evidence from Data on Robbery and Violent Theft." *Journal of Peace Research* 42, no. 1: 101–112.
- Osgood, D. Wayne. 2000. "Poisson-Based Regression Analysis of Aggregate Crime Rates." *Journal of Quantitative Criminology* 16, no. 1: 21–43.
- Ottaviano, Gianmarco, and Giovanni Peri. 2012. "Rethinking the Effect of Immigration on Wages." *Journal of the European Economic Association* 10, no. 1: 152–97.
- Peri, Giovanni. 2011. "Rethinking the Area Approach: Immigrants and the Labor Market in California." *Journal of International Economics* 84, no.1: 1–14.
- Pischke, Steve. 2007. *Lecture Notes on Measurement Error*.
econ.lse.ac.uk/staff/spischke/ec524/merr.pdf (accessed April 12, 2013).

- Sachsida, Adolfo; Mario Jorge Cardoso de Mendonça; Paulo R. A. Loureiro; and Maria Bernadete Sarmiento Gutierrez. 2010. "Inequality and Criminality Revisited: Further Evidence from Brazil." *Empirical Economics* 39, no. 1: 93–109.
- Santos, Marcelo Justus, and Ana Lúcia Kassouf. 2008. "Estudos Econômicos das Causas da Criminalidade no Brasil: Evidências e Controvérsias." *Revista Economia* 9, no. 2: 343–372.
- Saridakis, George. 2004. "Violent Crime in the United States of America: A Time-Series Analysis between 1960–2000." *European Journal of Law and Economics* 18, no. 2: 203–21.
- Scorzafave, Luiz Guilherme, and Milena Karla Soares. 2009. "Income Inequality and Pecuniary Crimes." *Economics Letters* 104, no. 1: 40–42.
- Spenkuch, Jorg L. 2014. "Understanding the Impact of Immigration on Crime." *American Law and Economic Review* 16, no. 1: 177–219.
- Souza, Paulo Renato. 2001. "Education and Development in Brazil, 1995–2000." *CEPAL Review* 73, no. 1: 65–80.
- Stoker, Thomas M. 2008. "Aggregation (Econometrics)." In *The New Palgrave Dictionary of Economics*, 2nd edition, eds. Steven N. Durlauf and Lawrence Blume. New York: Palgrave Macmillan, pp. 1–12.
- United Nations Office on Drugs and Crime (UNODC). 2009. *International Homicide Statistics*. New York: United Nations Publication.
- Vidal, Jean-Pierre. 1998. "The Effect of Emigration on Human Capital Formation." *Journal of Population Economics* 11, no. 4: 589–600.

World Bank. 2006. *Brazil – Crime, Violence and Economic Development in Brazil: Elements for Effective Public Policy*. Washington DC: World Bank.

Wu, Dongxu, and Zhongmin Wu. 2012. “Crime, Inequality and Unemployment in England and Wales.” *Applied Economics* 44, no. 29: 3765–3775.

Tables

Table 1. Basic statistics for 123 cities (agglomerations as defined in Mata et al. (2007)) for 1980, 1991 and 2000)

| Basic statistics for 123 agglomerations | | | | | | |
|--|-------------|-----------|-------------|-----------|-------------|-----------|
| | 1980 | | 1991 | | 2000 | |
| | Mean | SD | Mean | SD | Mean | SD |
| Homicides per 100,000 people | 23.56 | 16.990 | 31.26 | 23.25 | 40.59 | 29.48 |
| Homicides per 100,000 men | 48.33 | 35.91 | 64.44 | 48.8 | 83.6 | 61.21 |
| Police per 100,000 people | 375.78 | 239.63 | 492.91 | 282.49 | 287.26 | 169.22 |
| City average of median hourly wage | | | | | | |
| <i>High skill men</i> | 5.943 | 1.791 | 5.480 | 1.121 | 5.831 | 1.168 |
| <i>Low skill men</i> | 1.844 | 0.481 | 1.234 | 0.403 | 1.395 | 0.345 |
| <i>High skill women</i> | 2.71 | 0.79 | 2.88 | 0.78 | 3.42 | 0.72 |
| <i>Low skill women</i> | 1.03 | 0.24 | 0.77 | 0.25 | 0.95 | 0.22 |
| Male unemployment rate (%) | 19.14 | 6.33 | 18.17 | 5.19 | 28.26 | 5.94 |
| Average city size | 271,456 | 781,434 | 357,993 | 943,243 | 454,041 | 1,148,864 |
| <i>High skill men</i> | 9,894 | 33,137 | 14,260 | 47,149 | 19,157 | 59,462 |
| <i>Low skill men</i> | 106,045 | 303,146 | 146,247 | 373,477 | 177,147 | 430,659 |
| <i>High skill women</i> | 9,603 | 30,874 | 15,755 | 48,664 | 23,534 | 70,238 |
| <i>Low skill women</i> | 110,706 | 311,532 | 154,900 | 391,469 | 185,110 | 452,309 |
| <i>Percentage men</i> | 49.21 | 1.7 | 48.7 | 1.28 | 48.73 | 0.997 |
| <i>Percentage migrants</i> | | | 6.52 | 2.88 | 5.1 | 2.05 |
| <i>Percentage migrants (only men)</i> | | | 6.34 | 2.8 | 4.97 | 2.02 |
| <i>Ratio of high to low skill migrants</i> | | | 0.076 | 0.049 | 0.096 | 0.058 |

Notes: “High skill” is defined as having 12 years or more of education and “low skill” as less than 12 years of education.

“Migrants” are defined as people who moved into an agglomeration in the last five years, as recorded by the 1991 and

2000 Brazilian population censuses. SD stands for standard deviation. Total homicides data are sourced from DATASUS. Numbers for the police denote the total number of people working as members of a security force (excluding the army) from their respective census year. Unemployment is not reported consistently across census years; therefore, the non-participation rate in the labor market is reported as unemployment. City (agglomeration) size excludes recent migrants.

Table 2. Number of individuals incarcerated in Brazil in December 2008 by gender and age (Infopen, 2008)

| Incarcerated population by age group and gender (2008) | | | |
|---|---------|--------|---------|
| Age group (years) | Male | Female | Total |
| 18–24 | 113,635 | 5,686 | 119,321 |
| 25–29 | 96,058 | 5,160 | 101,218 |
| 30–34 | 63,475 | 3,903 | 67,378 |
| 35–45 | 53,924 | 4,135 | 58,059 |
| 46–60 | 20,800 | 1,729 | 22,529 |
| More than 60 | 3,174 | 154 | 3,328 |
| No information | 12,869 | 510 | 13,379 |
| Total | 371,884 | 21,604 | 393,488 |

Table 3. Number of individuals incarcerated in Brazil in December 2008 by gender and educational level (Infopen, 2008)

| Incarcerated population by educational attainment and gender (2008) | | | |
|--|---------|--------|---------|
| Educational attainment | Male | Female | Total |
| Illiterate | 27,192 | 1,24 | 28,432 |
| Literate | 44,582 | 2,422 | 47,004 |
| Incomplete basic education | 163,518 | 9,408 | 172,926 |
| Basic education completed | 46,476 | 2,786 | 49 |
| High school dropout | 39,212 | 2,489 | 41,701 |
| High school completed | 26,578 | 294 | 28,972 |
| Some university education | 3,301 | 417 | 4 |
| Bachelor's degree | 1,493 | 212 | 1,705 |
| Post-graduate degree | 61 | 7 | 68 |
| No information | 19,366 | 625 | 19,991 |
| Total | 371,884 | 21,604 | 393,488 |

4. Estimates of impact inequality (measured by the ratio of high to low skill wages) on homicides, assuming a Proportion for the probability of committing a crime

| regression on homicides | | | | | | |
|--|-----------------------------------|---------|----------|---------|----------|---------|
| | Change in log number of homicides | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ratio in high to low | 0.136*** | 0.046 | 0.043 | 0.029 | 0.06 | 0.023 |
| Age ratio | (0.045) | (0.049) | (0.047) | (0.048) | (0.043) | (0.049) |
| Ratio in young and low skill | -0.814*** | 0.288 | -0.162 | 0.203 | | |
| Average log hourly wage | (0.231) | (0.387) | (0.249) | (0.377) | | |
| Unemployment rate of young and low skill | 0.883*** | 0.0513 | 0.103 | 0.056 | | |
| Unemployment rate | (0.148) | (0.23) | (0.177) | (0.236) | | |
| Ratio of | 0.126 | 0.0717 | 0.044 | 0.014 | | |
| Force | (0.083) | (0.089) | (0.075) | (0.081) | | |
| Ratio of | | | 0.165 | 0.139 | 0.249 | 0.126 |
| Population | | | (0.415) | (0.401) | (0.383) | (0.395) |
| Age of recent young | | | 4.821*** | 3.733** | 4.873*** | 3.846** |
| Immigrants | | | (1.687) | (1.887) | (1.686) | (1.876) |
| Immigrants | No | Yes | No | Yes | No | Yes |
| Observations | 228 | 228 | 228 | 228 | 228 | 228 |

The model is in first difference, controlling for characteristics of cities, which do not change over time.

Robust standard errors are in parentheses.

Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 5. Percentage of high and low skill male and female workers in 3,214 rural municipalities working in different occupations between 1970 and 2000

| Occupational distribution of workers in rural areas in 1991 and 2000 | | | | |
|---|-------------------------|---------------------|-------------------------|----------------------|
| | 1991 | | 2000 | |
| | Low skill | High skill | Low skill | High skill |
| Number of rural out-migrants | 168.34 (551.06) | 26.2 (174.87) | 164.1 (557.71) | 29.58 (179.71) |
| Number of rural non-migrants | 5,757.37 (25,534.48) | 250.48 (5,654.9) | 6,588.44 (28,125.04) | 329.42 (6,114.44) |
| Occupations | | | | |
| <i>Administrative (%)</i> | 8.17 | 34.46 | 6.69 | 29.16 |
| <i>Technical or scientific (%)</i> | 1.87 | 36.76 | 2.95 | 39.37 |
| <i>Farming (%)</i> | 43.62 | 2.75 | 30.45 | 2.44 |
| <i>Mining (%)</i> | 1.43 | 0.18 | 0.60 | 0.08 |
| <i>Industry (%)</i> | 20.67 | 5.16 | 25.49 | 5.48 |
| <i>Commerce and trade (%)</i> | 7.99 | 7.19 | 9.87 | 8.99 |
| <i>Transport (%)</i> | 5.50 | 1.96 | 6.69 | 2.49 |
| <i>Services (%)</i> | 0.48 | 0.06 | 0.81 | 0.05 |
| <i>Domestic services (%)</i> | 3.36 | 2.40 | 5.24 | 3.06 |
| <i>Security and defense (%)</i> | 1.87 | 5.02 | 2.22 | 5.62 |
| <i>Other (%)</i> | 5.05 | 4.06 | 8.98 | 3.28 |
| Average rainfall (monthly average in cm) | 11.27 (4.19) | | 11.14 (4.34) | |
| Transport costs to São Paulo | 1,811 (1,437) | | 1,549 (1,126) | |
| Agricultural area (km ²) | 1,074 (4,716) | | 963 (4,537) | |
| Area (km ²) | 2,237 (13,093) | | | |

Source: 1991 and 2000 Brazilian population censuses.

Notes: The average and standard deviation (SD), in parentheses, of the number of migrants leaving rural areas appear for each year for the four years before the census, conditional on the municipality having rural out-migrants. The average rainfall and SD are monthly averages in centimeters for all months between 1986 and 1990, and between 1995 and 1999. Transport costs to São Paulo comprise an index centered around 1,000

and based on linear programming calculations by Castro (2002). The average agricultural (farming) area is for 1985 and 1995, sourced from the respective agricultural censuses. The area of the municipality is sourced from the 1970 census and is constant over time (the municipalities are comparable over time).

Table 6. Estimates of the impact of rainfall shocks and changes in transport costs from rural municipalities to São Paulo on out-migration from rural areas by group for 1986–1990 and 1995–1999

| Estimate of the impact of rainfall shocks and change in transport costs on rural out-migration | | |
|---|-----------------------|----------------------|
| | Log migrants | |
| | Low skill | High skill |
| Log lag natives | 0.7896*** (0.069) | 0.108 (0.063) |
| Log agricultural area (ha) | -0.046 (0.032) | -0.172** (0.081) |
| Transport costs to São Paulo | 0.079 (0.095) | -0.849** (0.3496) |
| Average monthly rainfall (mm) | -0.0052** (0.0024) | -0.016 (0.011) |
| Previous year's average monthly rainfall | -0.0052** (0.0024) | -0.016 (0.011) |
| Average monthly rainfall in semi-arid area | 0.0201*** (0.0055) | 0.1301*** (0.037) |
| Previous year's average monthly rainfall in semi-arid area | 0.032*** (0.0061) | 0.097** (0.037) |
| Observations | 25 712 | 22 338 |
| Fraction of observations censored (%) | 2.57 | 72.01 |
| Municipality fixed effects | Yes | Yes |
| Year dummy variables | Yes | Yes |
| Number of municipalities | 3,214 | 3,21 |
| χ -squared test on parameters (<i>p</i> -value) | 44.79 (0) | 26.43 (0) |

Notes: Cluster-robust standard errors are in parentheses.

* Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 7. Estimates of the importance of distance between the origin (3,214 municipalities) and destination (123 cities) in the choice of destination, pooling each group

| Estimates for distance as an explanation for migration location decision | | | | |
|---|------------------------------|-----------------------|-------------------------|------------------------|
| | Percent rural–urban migrants | | | |
| | (1) | (2) | (3) | (4) |
| Log distance | -0.017** (0.00018) | -0.021** (0.00025) | -0.022** (0.00027) | -0.022** (0.00026) |
| Log distance (for high skill) | | 0.0101** (0.00034) | 0.013** (0.00043) | 0.011** (0.0004003) |
| (Log) People living in rural area X | | | 0.00103** (0.000011) | 0.0014** (0.000024) |
| (Log) People living in urban area | | | | |
| Log number of people living in rural area | | | | -0.0082** (0.00024) |
| Log number of people living in urban area | | | | -0.0011** (0.00014) |
| Dummy for high skill | Yes | Yes | Yes | Yes |
| Municipality fixed effects | Yes | Yes | Yes | Yes |
| Observations | 787,815 | 787,815 | 787,815 | 787,815 |
| Number of rural municipalities | 3,214 | 3,214 | 3,214 | 3,214 |
| <i>R</i> -squared | 0.04 | 0.05 | 0.07 | 0.07 |

Notes: Cluster-robust standard errors are in parentheses.

* Significant at 10%, ** significant at 5%, *** and significant at 1%.

Table 8. First stage of instrumental variables regression for changes in high to low skill wage ratio

| OLS first-stage regression for inequality | | | | | | | | |
|--|---|----------|------------|-----------|--|-----------|-----------|-----------|
| | Percentage of recent young and low skill migrants | | | | Change in high to low skill wage ratio | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Predicted changes in the | 0.00817*** | -0.00431 | 0.00793*** | -0.004 | 0.175*** | -0.859*** | 0.033 | -0.769*** |
| ln of recent migrants | (0.0013) | (0.0096) | (0.0013) | (0.00904) | (0.038) | (0.141) | (0.028) | (0.144) |
| Predicted change in | -17.12*** | -9.451* | -10.24** | -8.15* | -650.4*** | 387.3*** | 216.6** | 349.6*** |
| high to low skill migrants | (4.096) | (4.83) | (4.14) | (4.56) | (111.5) | (82.52) | (83.29) | (81.39) |
| Growth rate of | 0.115*** | 0.107*** | 0.112*** | 0.109*** | 1.633*** | 0.56 | 0.61 | 0.43 |
| city population | (0.018) | (0.017) | (0.017) | (0.017) | (0.467) | (0.41) | (0.459) | (0.451) |
| Change in young and low skill | | | 0.0174* | 0.033** | | | -3.146*** | -2.24*** |
| males' average log hourly wage | | | (0.0101) | (0.014) | | | (0.403) | (0.568) |
| Growth rate of young and low | | | -0.0229* | -0.017 | | | -0.581 | -0.227 |
| skill males' unemployment rate | | | (0.012) | (0.015) | | | (0.45) | (0.527) |
| Growth rate of | | | 0.0117*** | 0.0101** | | | 0.021 | -0.07 |
| police force | | | (0.0041) | (0.00397) | | | (0.153) | (0.155) |
| Year dummies | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 228 | 228 | 228 | 228 | 228 | 228 | 228 | 228 |
| R-squared | 0.859 | 0.362 | 0.871 | 0.397 | 0.374 | 0.508 | 0.661 | 0.541 |
| F-test | 21.63 | 5.445 | 21.48 | 4.155 | 104.7 | 21.08 | 9.987 | 18.33 |

Notes: As instruments, I use the predicted log number of total recent migrants and the predicted change in high to low skill migrants (see Section IV).

Standard errors are in parentheses and are only efficient for homoskedasticity.

* Significant at 10%, ** significant at 5%, and *** significant at 1%.

Table 9. Estimates of the impact of inequality (measured by the ratio of high to low skill wages) on homicides, for the Poisson model

| Second-stage of IV regression on homicide (GMM) | | | | | | | | |
|--|---------------------------------------|-------------------|--------------------|--------------------|------------------|--------------------|-------------------|--------------------|
| | Change in the log number of homicides | | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Change in high to low skill wage ratio | -0.123 (0.165) | 0.372** (0.18) | 0.579** (0.229) | 0.425** (0.213) | 0.267 (0.223) | 0.434** (0.183) | | |
| Change in 75 th percentile wage to 25 th percentile wage ratio | | | | | | | 5.233 (10.21) | 2.36*** (0.897) |
| Percentage of recent young and low skill migrants | 8.546** (3.67) | -0.763 (6.36) | 1.936 (3.09) | -2.404 (8.4) | 2.673 (3.75) | 6.805 (13.41) | 17.57 (24.17) | 3.87 (15.22) |
| Growth rate of city population | -0.046 (0.535) | 0.311 (0.870) | -0.029 (0.653) | 0.565 (1.126) | 0.454 (0.545) | 0.086 (1.002) | 1.349 (2.024) | 1.306 (0.94) |
| Change in young and low skill males' average log hourly wage | | | 0.014 (0.299) | -0.101 (0.442) | 0.636 (0.704) | 1.295** (0.623) | 7.346 (14.66) | 3.39*** (1.273) |
| Growth rate of young and low skill males' unemployment rate | | | 1.482* (0.87) | 1.308* (0.692) | 0.002 (0.234) | 0.245 (0.469) | -0.862 (1.759) | -0.505 (0.403) |
| Growth rate of police force | | | -0.009 (0.111) | 0.017 (0.122) | 0.046 (0.094) | -0.009 (0.124) | 0.131 (0.254) | 0.113 (0.117) |
| Year dummies | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 228 | 228 | 228 | 228 | 227 | 227 | 227 | 227 |
| Kleibergen–Paap rk LM-statistic (<i>p</i> -value) | 0 | 0.0018 | 0.001 | 0.0079 | 0.0003 | 0.0074 | 0.563 | 0.0642 |
| Stock–Wright LM <i>S</i> -statistic (<i>p</i> -value) | 0.0521 | 0.0165 | 0.0203 | 0.0111 | 0.2620 | 0.0144 | 0.262 | 0.0144 |

Notes: All regressions are estimated using generalized method of moments (GMM). The sample is restricted to young and low skill men. The model is in the first difference, controlling for characteristics of cities, which do not change over time. In columns (1)–(4) and (7)–(8), the sample includes only young men (15–45 years), while in columns (5)–(6), the sample includes both men and women

between the ages of 15 and 65 years.

Cluster-robust standard errors are in parentheses.

* Significant at 10%, ** significant at 5%, and *** significant at 1%

Figures

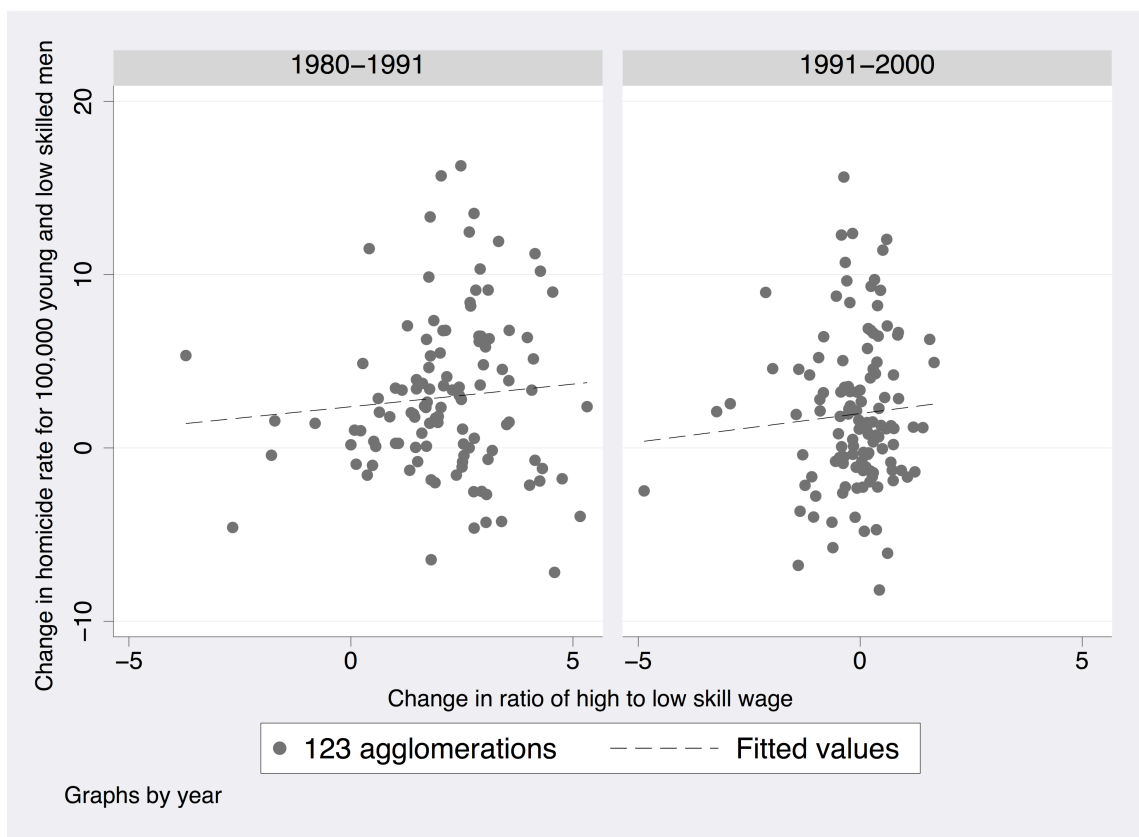


Figure 1. Relationship between the change in the homicide rate by 10,000 young and low skill men in each agglomeration and the change in the ratio of high to low skill wage (a proxy for income inequality).

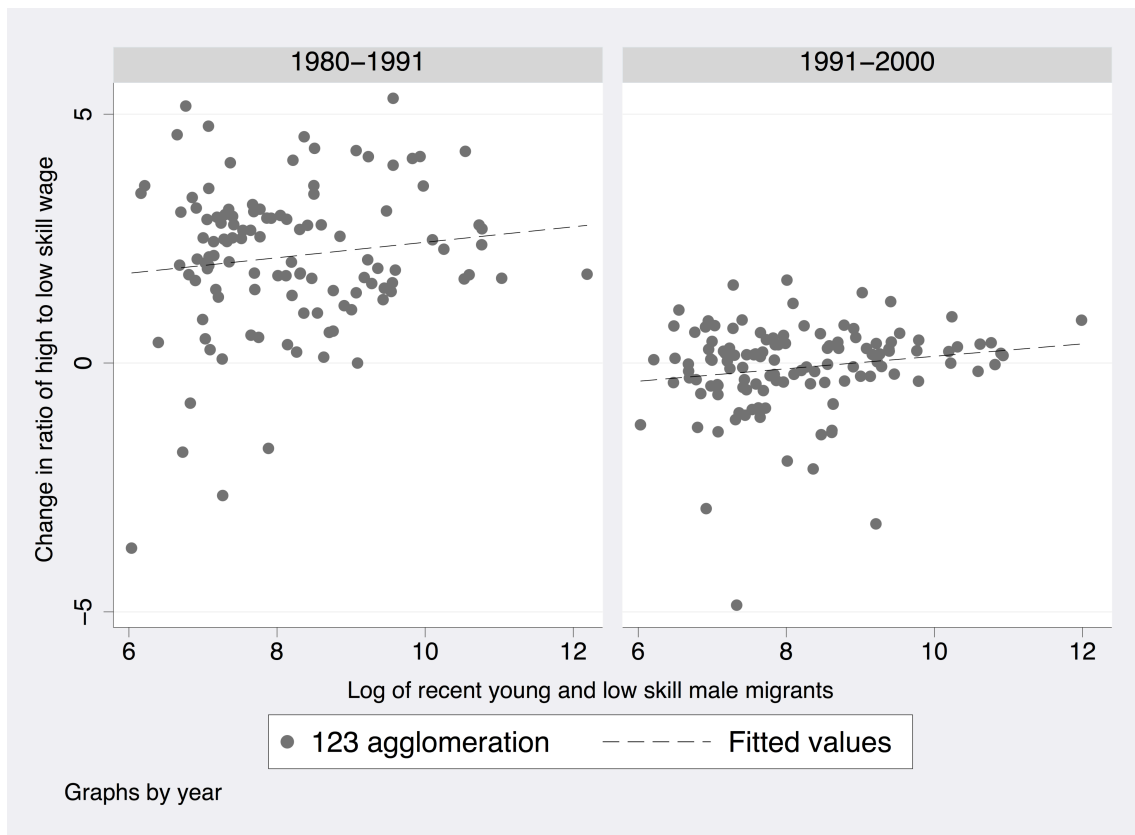


Figure 2. Relationship between change in the ratio of high to low skill wages (a proxy for income inequality) in each agglomeration and the log of recent young and low skill male migrants

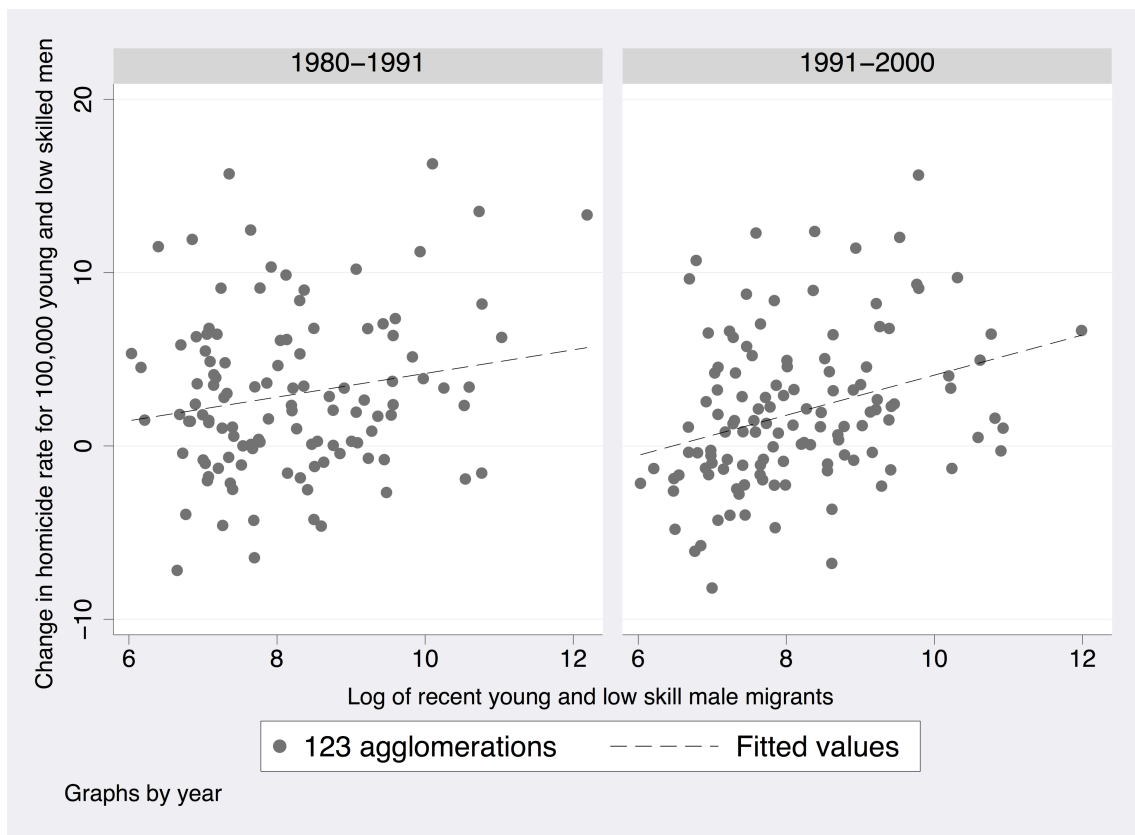


Figure 3. Relationship between the change in the homicide rate by 100,000 young and low skill males in each agglomeration and the log of recent young and low skill male migrants.