The University of Adelaide
School of Economics

Research Paper No. 2011-02
January 2011

Cultural Diversity and Economic Growth: Evidence from the US during the Age of Mass Migration

Philipp Ager and Markus Brückner

Cultural Diversity and Economic Growth: Evidence from the US during the Age of Mass Migration.*

Philipp AGER and Markus BRÜCKNER

December 2010

Abstract: We exploit the large inflow of immigrants to the US during the 1870-1920 period to examine the effects that changes in the cultural composition of the population of US counties had on output growth. We construct measures of fractionalization and polarization to distinguish between the different effects of cultural diversity. Our main finding is that increases in cultural fractionalization significantly increased output, while increases in cultural polarization significantly decreased output. We address the issue of identifying the causal effect of cultural diversity on output growth using the supply-push component of immigrant inflows as an instrumental variable.

Keywords: Cultural Diversity, Economic Growth, Historical Development, Immigration.

JEL: O1, Z1

*Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. University of Adelaide, 10 Pulteney Street, 5000 Adelaide. Contact e-mail: philipp.ager@upf.edu; markus.bruckner@adelaide.edu.au. We are grateful to Marta Reynal-Querol for helpful comments and suggestions. We also thank Leah Platt Boustan, Antonio Ciccone, Joan Esteban, Libertad González, Ian W. McLean, Thijs van Rens, Dietrich Vollrath and seminar participants of the UPF Labour/Public/Development Workshop for further comments and discussion. We thank Robert A. Margo and Jeremy Atack for providing us with their historical dataset on US county railroads. All remaining errors are our own.

1 Introduction

During the 1850-1920 period the US experienced a mass inflow of immigrants – more than 30 million people migrated from Europe to the US – with the average annual immigration inflow rate measuring about 0.75 percent of the total US population (Hatton and Williamson, 1998). We exploit in this paper this historically unique inflow of immigrants to the US to study the effects that changes in the cultural composition of the US population had on output growth. Because immigrants came from different European countries, the mass immigration wave not only affected the overall share of foreign-born in the US. It also affected significantly the diversity of the working and voting-age population.[1]

A fairly large and by now well-established literature has investigated the effects of cultural diversity on economic growth in the cross-section of countries.[2] Our aim is to contribute to this literature in two main dimensions. First, our empirical analysis is based on comparing how within-county changes in cultural fragmentation affect within-county changes in output. Because we use in the estimation exclusively the within-county variation of the data, our empirical analysis circumvents an often made critique of the cross-country growth literature that in the cross-section there are many difficult-to-measure omitted variables, such as history and geography, that affect both economic growth and the cultural fragmentation of the population. Second, we address the important issue of identifying the causal effect that cultural diversity has on economic growth by using the supply-push component of immigrant inflows as an instrumental variable. The supply-push component of immigrant inflows is a widely used instrumental variable in the labor economics literature (see e.g. Card and DiNardo, 2000; Card, 2001, 2009; Saiz, 2003; Ottaviano and Peri, 2005, 2006) and has not been used before to examine at the US county level the effects that cultural diversity had on output growth during the age of mass migration.

On the theoretical front, the cultural economics literature has suggested several potentially countervailing channels through which cultural diversity affects output growth. Cultural diversity can have a positive effect on output growth if a more diverse working-age population is associated with a greater variety of skills that in turn enable the production of a greater variety of goods and services (Alesina and La Ferrera, 2005). A more diverse, and in particular, a more polarized population can however also have a detrimental effect on output growth if it is associated with increased social conflict and a reduction in the quality and quantity of public good provision (Alesina et al., 1999; Montalvo and Reynal-Querol, 2005a,b). We examine these countervailing effects by constructing measures of cultural fractionalization and cultural polarization. While our measure of cultural fractionalization is linearly increasing in the number of groups, our measure of cultural polarization is maximized when two groups are

[1] See for example, Cohn (2010), Hatton and Williamson (1998), or Kim (2007) and the references therein for further information on the history of immigration to the US.

[2] See for example, Mauro (1995), Easterly and Levine (1997), Alesina et al. (2003), or Montalvo and Reynal-Querol (2003, 2005a).

of equal size. Our polarization measure is therefore closely related to Horowitz's (1985) statement that conflicts are more likely in societies where a large ethnic minority faces an ethnic majority. Conflict models such as those in Lane and Tornell (1996, 1999), Esteban and Ray (1999, 2010), and Montalvo and Reynal-Querol (2005a) formalize the non-monotonic relationship between diversity and potential conflict emphasized by Horowitz (1985) and predict that social tensions are greatest when there are two equally powerful groups that contest for resources. We expect our polarization measure therefore to capture relative to the fractionalization measure more the social conflict effect than the production diversity effect.

Our first main finding is that increases in the cultural fractionalization of US counties during the 1870-1920 period led to significant increases in output per capita. State-fixed effects estimates yield that on average a 1 percentage point increase in cultural fractionalization increased output per capita by up to 2.3 percent. This is a sizable effect. Consistent with the results in Alesina and La Ferrera (2005) we also find that the positive effect of fractionalization on output growth is particularly large in counties with above-median GDP per capita. In these counties a 1 percentage point increase in cultural fractionalization increased output per capita by up to 4.2 percent.

As an intermediate channel to link our first main result more closely to the cultural economics literature – where a key argument for a positive effect of cultural fractionalization on output per capita is increased product variety (product diversity channel) – we show that increases in the cultural fractionalization of counties were associated with significant increases in the occupational diversity of workers.[3] We construct a fractionalization index of worker occupation and show that changes in this index were significantly positively affected by changes in the cultural fractionalization of counties. Consistent with the product diversity channel, we also find that increases in cultural fractionalization led to a greater variety of agricultural goods.

Our second main finding is that increases in cultural polarization had a significant negative effect on output per capita. A state-fixed effects estimate yields that a 1 percentage point increase in cultural polarization decreased output per capita by around 2.5 percent. We provide an explanation for this negative effect of cultural polarization on output by documenting that increases in the cultural polarization of counties were associated with a significant increase in the fraction of people employed in jobs related to crime delicts. We also document that increases in cultural polarization led to a significant increase in the number of riots. Our finding of a significant positive effect of cultural polarization on the fraction of people employed in crime related jobs and a significant positive effect of cultural polarization on the number of riots is consistent with the conflict literature that has emphasized the conflict promoting effects of cultural polarization (e.g. Montalvo and Reynal-Querol, 2003, 2005a,b, 2008; Esteban and Ray, 1999, 2010 and Esteban et al., 2010).

[3] Our finding is in line with Kim (2007), who provides descriptive evidence that immigration and the diversity of the workforce are positively related.

An additional contribution of our paper is that our state-fixed effects estimates reveal a significant reversion of counties' cultural fractionalization and cultural polarization index to the mean. Statistically, this mean reversion follows from the bounded nature of the polarization and fractionalization index on the unit interval. The significant mean-reversion in the cultural fractionalization and polarization index implies that, in a reduced form regression of economic growth on initial fractionalization and polarization the estimated coefficients reflect the opposite of the true effect that a within-county change in cultural fractionalization and polarization has on within-county output changes. This is because counties with an initially high level of cultural fractionalization and polarization experienced subsequently smaller changes than counties with initially low levels. Hence, using initial values of fractionalization and polarization to examine how a change in the cultural fragmentation of the population induces a change in output per capita requires to deal correctly with the mean-reverting time-series nature of the cultural fractionalization and polarization indices. We show that this can be done by using an IV approach where the change in fragmentation is instrumented by the initial level of fragmentation. Via the first-stage regression the IV approach fully takes into account the mean-reversion in cultural fractionalization and polarization. Importantly, we find that our main results continue to hold when using this alternative instrumental variables approach which exploits the cross-county variation in initial conditions.

Methodologically, our empirical analysis is most closely related to Ottaviano and Peri's (2006) work on the economic value of cultural diversity in US cities. Ottaviano and Peri tackle endogeneity problems by using, as we do, the identification strategy from Card (2001) and Saiz (2003) to construct the predicted change of the number of immigrants from each country in each US city between 1970 to 1990 based on immigration network effects (i.e. the supply-push component). Consistent with our results, Ottaviano and Peri find that increases in the cultural fractionalization of cities has a positive effect on productivity.[4]

Our empirical analysis differs from the work of Ottaviano and Peri in two main aspects. First, Ottaviano and Peri focus on cultural fractionalization only and do not examine the effects of cultural polarization. Our estimates show that increases in cultural polarization have a significant negative output effect while increases in cultural fractionalization have a significant positive output effect. The empirical findings in our paper therefore highlight the importance of distinguishing between cultural fractionalization and cultural polarization. They echo the theoretical literature that has emphasized different channels through which cultural diversity can affect economic activity. In Appendix A1 we also present anecdotal evidence on these countervailing effects of cultural polarization and cultural fractionalization to further illustrate our empirical findings.

[4]In complementary work, Ottaviano and Peri (2005) show that employment density and wages of US-born workers are significantly higher in US cities with a higher linguistic fractionalization. More recently, Sparber (2009, 2010) shows with decennial panel data that racial fractionalization has a significant positive effect on wages in most of the US industries and cities for the 1980-2000 period. Peri (2010) shows for the 1970-2006 period that immigration had a significant positive effect on total factor productivity in US states.

A second key difference between Ottaviano and Peri and our empirical analysis is the time period covered. Ottaviano and Peri focus on the 1970-1990 period while we focus on the period from 1870-1920. There are several reasons why focusing on the 1870-1920 period has advantages. First, in contrast to the 1970-1990 period, the 1870-1920 period was a period of free immigration. Illegal immigration was therefore not an issue due to the very liberal US immigration policy before 1920. Second, immigration during the 1870-1920 period changed significantly the demographic structure of the US population. Between 1870-1920, the annual inflow rate of immigrants peaked in some years around 1.5 percent, a value which was never reached again in the 20th century (see Figure 1). Finally, from a historical perspective, the period of mass migration constituted the first large inflow of immigrants to the US.

The remainder of our paper is organized as follows. Section 2 provides a brief overview and discussion of the age of mass migration. Section 3 describes the construction of our cultural fractionalization and cultural polarization index. Section 4 discusses the estimation strategy. Section 5 presents our main results. Section 6 examines the issues of group aggregation and the measurability of inter-group distances. The last section concludes.

2 The Age of Mass Migration in the US: A Natural Source of Cultural Variation

During the age of mass migration, which is commonly referred to as the period between 1850 to World War I, about 55 million Europeans emigrated to North and South America and Australasia. The US received about three-fifth of the 55 million European immigrants, and thus the bulk of the large immigration stream (Hatton and Williamson, 1998). In general, between 1850 and 1920, emigrants from non-European countries (like China, Japan and Mexico) constituted only a minor part of the overall immigration wave to the US (see e.g. Cohn, 2010).

One important and key aspect of the age of mass migration is that it was a period of free immigration. Despite the Chinese exclusion act in 1882, immigration policy remained liberal for the overwhelming European part of immigrants until the introduction of a literacy test in 1917 (Immigration Act of 1917) and the establishment of immigration quotas in 1921 (Emergency Quota Act of 1921).[5] The 1850-1920 immigration inflow was therefore a natural source of variation in cultural diversity. It stands in contrast to the post-1920 US immigration inflows, which were strongly affected by US immigration policy.

We focus in our empirical analysis on the period 1870 to 1920, rather than on the whole 1850 to 1920 period for two main reasons. First, we wanted to leave out the 1861-1865 civil war. The 1861-1865 civil war was an atypically large, negative shock to the US economy that could have affected both the cultural

[5] See Goldin (1994) for an excellent analysis of the political economy of immigration restrictions in the United States, 1890 to 1921.

diversity and development of US counties differently.[6] Second, we take 1920 as the terminal point because at the turn of the 1920s there was a significant change in US immigration policy. As Goldin (1994, p. 223) notes: "With the passage of the Emergency Quota Act in May 1921 the era of open immigration to the United States came to an abrupt end".

Figure 1 shows that during the 1870-1920 period the annual inflow rate of immigrants was about 0.75 percent of the total US population. This is a fairly large inflow rate, in particular when compared to the post-1920 period where among other factors immigration restrictions significantly reduced the inflow rate of immigrants to about 0.20 percent.

Table 1 presents descriptive statistics of the share of foreign-born in the total US population during the 1870-1920 period. On average, more than 10 percent of the US population were foreign-born. The West had the largest share of foreign-born (19 percent), followed by the Midwest (15 percent), the Northeast (15 percent), and the South (2 percent). The descriptive statistics show also that not only is there substantial variation in the foreign-born share across US counties – the across-county standard deviation of the foreign born share is 0.11. The share of foreign-born changed also significantly within counties (the within-county standard deviation is 0.04). Because the population in the US more than doubled between 1870-1920, in order to keep the share of foreign-born constant at 10 percent the absolute number of immigrants had to increase (see e.g. Barde et al., 2006).

It is also interesting to note that during the age of mass migration the geographic origin of the immigrants varied substantially over time. Around 1870, most of the immigrants were from Ireland, Germany, Great Britain, and Scandinavia. These countries remained until the end of the 19th century very important emigration nations. At the end of the 19th century, however, the bulk of immigrants came from Eastern and Southern Europe including countries like Italy, the Austro-Hungarian Empire, Poland, Russia, Spain or Portugal and replaced Scandinavia, Germany, Ireland and Great Britain as the most important immigration nations in the US (see e.g. Barde et al. 2006, Cohn, 2010, or Hatton and Williamson, 1998). This geographical shift of immigration from Western and Northern Europe to Southern and Eastern Europe resulted in a remarkable variation of the cultural composition of the US county population between 1870 and 1920. We explain in the next section how we calculated indices of cultural fractionalization and cultural polarization to capture these substantial changes in the cultural diversity of US counties.

[6] Furthermore, we start our empirical analysis in 1870 because the US Census collected systematical information about the foreign-born's country of origin from 1870 onwards only (see the Data Appendix A3 for further information).

3 Two Different Concepts of Cultural Diversity: Fractionalization and Polarization

We use the country of origin defined by the US Census to measure different cultural groups and to construct our indices of cultural diversity. In contrast to ethno-linguistic or religious fragmentation indices that distinguish among different races, linguistic, or religious characteristics of individuals, our indices are based on the country of origin of individuals and this has the key advantage that our indices are based on well-defined units. We construct an index of cultural fractionalization as

$$FRAC_{cs} = 1 - \sum_{i=1}^{N} \pi_{i,cs}^2,$$

where $\pi_{i,cs}$ is the county population share of group i living in county c of state s. The N groups are: US-born whites, Afro-Americans, foreign-born from the Austro-Hungarian Empire (Austria, Bohemia and Hungary), the Benelux (Holland, Luxembourg, Belgium), Canada, Central and South America, Scandinavia (Denmark, Norway, Sweden and Finland), Eastern Europe, France, Germany, Greece, Ireland, Italy, Poland, Portugal, Spain, Switzerland, and the United Kingdom (England, Scotland and Wales).[7] Conceptually, the fractionalization index captures the probability that two randomly selected individuals are from different cultural groups. In general, the fractionalization index is one minus the Herfindahl-Hirschman index and has the property that it increases monotonically in the number of groups. Due to the common use of the Herfindahl-Hirschman index in the industrial organization literature, the fractionalization index should be well suited to test whether there is a benefit from cultural diversity inherent in the production process or in the supply of different consumption goods.

To capture the degree of social tensions, we construct a measure of cultural polarization as

$$POL_{cs} \quad = \quad 1 - \sum_{i=1}^{N} \left(\frac{1/2 \; - \; \pi_{i,cs}}{1/2} \right)^2 \pi_{i,cs}.$$

Our polarization measure follows Reynal-Querol (2002) and Montalvo and Reynal-Querol (2005b) and captures formally how far the distribution of the cultural groups is from the bipolar distribution.[8] The measure is therefore maximized when there are two groups which are of equal size. Note that the polarization index that we use is based on the binary criteria of "belonging" or "not belonging" to a particular cultural group. Using such a discrete metric (belonging/not

[7] See also the Data Appendix A3 for further details.

[8] For further details on the theoretical properties of a discrete polarization measure see Montalvo and Reynal-Querol (2008).

belonging) implies that the distance across ethnic groups is the same. Hence, what matters for our index of polarization is only the size of each group and not the distance between groups.[9] Given the work by the conflict literature on polarization the polarization index seems a proper tool to analyze the economic consequences of potential conflicts among cultural groups in the US.

Figures 2A and 2B show scatter plots of the polarization and fractionalization index. Figure 2A shows a scatter plot for the 1870 period and Figure 2B shows a scatter plot for 1920 period. Both scatter plots show an inverted U-shaped relationship. For low values of fractionalization the relationship is positive, for intermediate values the relationship is zero, and for high values of fractionalization the relationship is negative. Hence, the correlation between the polarization and fractionalization index is low when the cultural heterogeneity is high.[10]

To show that there was also significant variation in fractionalization and polarization over time, Table 2 provides summary statistics of the fractionalization and polarization index for the 1870 and 1920 period. Interestingly, both fractionalization and polarization declined in the US on average. But the change in fractionalization and polarization differed substantially across regions. In the Northeast both polarization and fractionalization increased, while in the West, the Mid-West, and the South polarization and fractionalization decreased.

4 Estimation Strategy

We estimate the effects that cultural fractionalization and polarization have on output per capita using a first-difference approach which relates the 1870-1920 change in output per capita ΔY_{cs} of US county c in state s to the 1870-1920 change in fractionalization $\Delta FRAC_{cs}$ and polarization ΔPOL_{cs}:

$$\Delta Y_{cs} = \alpha_s + \beta \Delta FRAC_{cs} + \gamma \Delta POL_{cs} + \Gamma' X_{cs,1870} + \Delta u_{cs}. \quad (1)$$

As a baseline we estimate equation (1) by least squares. The state fixed effects α_s capture factors that vary over time at the state-level (such as for example economic and social policy choices made by the state government) and affect both output per capita and the cultural diversity of the population. Note that because we estimate equation (1) in first differences, time-invariant factors (such as for example geography) that affect both output per capita and cultural diversity are differenced out.

On theoretical grounds, we cannot rule out that changes in output per capita of counties had no systematic effects on their cultural diversity. For example,

[9] In Section 6.2 we also discuss results for an alternative index that attempts to capture distances between groups.

[10] In Appendix Figure 1 we provide further intuition on the relationship between the fractionalization and polarization index by simulating data and plotting the relationship between the fractionalization and polarization index for different group numbers (see also Appendix A2 for further details).

counties that experienced a large increase in output per capita may appear particularly attractive to immigrants, causing an inflow of immigrants. It is also possible that due to demand effects counties which grew a lot demanded a more diverse workforce. In this case, least-squares estimation of equation (1) will produce biased estimates.

To address the concern that the least-squares estimates are biased due to a reverse causal effect of economic growth on cultural diversity we use instrumental variables techniques. Our first instrumental variables approach follows the labor economics literature and uses the so-called supply-push inflow of immigrants as an instrumental variable. For the 1870-1920 period, we compute the supply-push inflow of immigrants from European country i to US county c as

$$SP_{i,cs} = \theta_i^{7020} \pi_{i,cs}^{1870},$$

where $\pi_{i,cs}^{1870}$ is the share of foreign-born from European country i living in year 1870 in US county c in state s and $\theta_i^{7020} = \left(\frac{\pi_i^{1920} - \pi_i^{1870}}{\pi_i^{1870}} \right)$ denotes the nation-wide growth rate of the foreign-born share from European country i between 1870-1920. The identification strategy therefore exploits that newly arriving immigrants tend to settle in regions where previous immigrants from the same country already live (see e.g. Bartel, 1989, Munshi, 2003 or Boustan, 2010). Because we use the initial, that is, the 1870 distribution of immigrants across the US counties to predict the inflow of immigrants during the 1870-1920 period the supply-push variable is exogenous to demand shocks that occur over the 1870-1920 period in US counties. The imputed foreign-born share, $\widehat{\pi}_{i,cs}^{1920}$, is calculated as $\widehat{\pi}_{i,cs}^{1920} = (1 + \theta_i^{7020}) \pi_{i,cs}^{1870}$. Because the fractionalization and polarization index are functions of $\pi_{i,cs}^2$ and $\pi_{i,cs}^3$ we use the change in the quadratic and cubic terms of the imputed foreign-born shares as instrumental variables for the change in the fractionalization and polarization index.

Our second instrumental variables approach exploits that there is significant reversion of the fractionalization and polarization index to the mean. The significant mean-reversion in the fractionalization and polarization index implies that we can use the initial 1870 values as instrumental variables to predict the 1870-1920 change. Because the initial values are predetermined the instruments are plausibly exogenous to future changes in output per capita.

We note that even though our predetermined instruments are well suited to address reverse causality bias, they may not be so well suited to address omitted variables bias. We therefore made an effort to address omitted variables bias in our estimation strategy as much as possible by including in the regression state fixed effects as well as a rich set of county-specific initial control variables $X_{cs,1870}$. The set of county-specific initial control variables includes amongst others the (1870) level of GDP per capita, the Gini coefficient on land inequality, the manufacturing share, and an indicator variable for railway access. Including these initial control variables on the right-hand side of the estimating equation should increase the likelihood that our predetermined instruments affect economic growth over the 1870-1920 period only through their effect on the

1870-1920 change in counties' fractionalization and polarization. We examine in our empirical analysis also formally the validity of our instrumental variables by using the Hansen test of the overidentifying restrictions.

5 Main Results

Table 3 presents our baseline estimates of the average effect that cultural fractionalization and cultural polarization have on output growth. Column (1) reports the state-fixed effects estimates for regressing the 1870-1920 log-change of output per capita on the 1870-1920 change in the polarization index. The result is a negative coefficient on the polarization index that is statistically significant at the 10 percent level. In column (2) we repeat the regression for the cultural fractionalization index and this produces a negative but statistically insignificant coefficient.

Results change somewhat when we include both, the fractionalization and the polarization index on the right-hand side of the estimating equation. Column (3) shows that in this case both the fractionalization and the polarization index are statistically significant at the 10 percent level. The estimated coefficient (s.e.) on the polarization index is -1.08 (0.41) and the estimated coefficient (s.e.) on the fractionalization index is 1.36 (0.72). Hence, the effect of polarization on economic growth remains significantly negative, while the growth effect of cultural fractionalization becomes significantly positive.[11]

In column (4) we introduce control variables that capture cross-county differences in initial conditions, such as initial output per capita, the Gini coefficient of farm size distribution, the manufacturing share, the population size, rail access, the share of foreign-born, and the population share of Afro-Americans. Controlling for these variables substantially improves the explanatory power of the econometric model: the adjusted R-squared increases from 0.27 in column (3) to 0.54 in column (4). And the significance of the estimated effect of fractionalization and polarization on output growth increases so that both point estimates on the fractionalization and polarization index are individually significant at the 1 percent level.

To correct for a possible endogeneity bias of our least-squares estimates, we report in column (5) IV estimates that use the supply-push component of immigrant inflows as an instrumental variable. The main result is that there continues to be a significant negative growth effect of cultural polarization and a significant positive growth effect of cultural fractionalization when using the supply-push instrumental variables strategy. In terms of instrument quality, the supply-push strategy yields a reasonable first-stage F-statistic of 11.01 and the Anderson-Rubin test rejects at the 1 percent level that the second-stage coef-

[11] We note here that it is important to include both fractionalization and polarization jointly in the regression because these variables are positively correlated (see Section 3). Including fractionalization (polarization) individually implies that the estimate suffers from a negative (positive) omitted variables bias since in this case the right-hand side regressor is correlated with the error term.

ficients are equal to zero. Quantitatively, the two-stage least squares estimates imply that a 1 percentage point increase in cultural polarization significantly reduced output per capita by over 1.7 percent. A 1 percentage point increase in cultural fractionalization significantly increased output per capita by around 1.9 percent.

In column (6) we show that similar results are obtained when using instead of the supply-push immigrant inflow instrument, counties' initial level of fractionalization and polarization as an instrumental variable for the 1870-1920 period change. This instrumental variables strategy, which exploits the significant reversion of the fractionalization and polarization measure to the mean produces 2SLS coefficients on the polarization and fractionalization index of -2.5 and 2.3 respectively. Both 2SLS coefficients are individually significant at the 1 percent level at least. The joint first-stage F-statistic of the excluded instruments is 42.81 and thus easily exceeds the critical values tabulated in Stock and Yogo (2005) for instruments to be declared as weak.

In order to formally test the validity of the instrumental variables in the second-stage growth equation, we report in column (7) two-stage least squares estimates that use both the supply-push and the initial fractionalization and polarization index as instrumental variables. The Hansen test on the overidentifying restriction produces a p-value of 0.79. Hence, the Hansen test does not reject the null hypothesis that the instrumental variables are uncorrelated with the second-stage error term.

As a further check on the robustness of the estimated average marginal effect that polarization and fractionalization has on output growth, Table 4 presents least squares and instrumental variables estimates for different time periods. Our baseline regressions are for the period from 1870 to 1920 because our aim is to examine the long-run effect that changes in the cultural diversity of counties had on output growth during the age of mass migration. To check whether our results also hold when shortening the time span, Panel A of Table 4 presents least squares and instrumental variables estimates for the 1880-1920 period, Panel B presents estimates for the 1890-1920 period, and Panel C presents estimates for the 1900-1920 period.[12] The main result is that for these alternative time periods cultural fractionalization has a positive effect on output growth and cultural polarization has a negative effect. Statistically, the estimates are significant at conventional confidence levels for the majority of the alternative time periods.[13] In Appendix Table 1, we also show that there continues to be a negative average effect of polarization on output growth and a positive average effect of fractionalization when excluding the Southern states (where output growth and immigration on average was low and polarization was high) and in Appendix Table 2 we show that our results are robust to controlling for a nonlinear effect of fractionalization.

[12] Because no manufacturing data are available for the 1910 US Census at the county level, we are not able to examine the effect for the 1910-1920 period.

[13] Least squares and instrumental variables estimates for the periods 1870-1880, 1870-1890 and 1870-1900 yielded qualitatively similar results, which are available from the authors upon request.

The economic history literature has often used as a proxy for economic progress and development the growth rate of the share of people living in urban regions. Although historically changes in output per capita and urbanization might have been highly correlated, Table 5 shows that the results are actually not as strong when we use instead of the output per capita growth rate the change in the share of people living in urban regions. The estimated coefficients still show a positive effect of fractionalization and in some specifications a negative effect of polarization. But, statistically the estimates are mostly insignificant. Given that output per capita can be much more readily related to macroeconomic models with an aggregate production function than urbanization rates, we are not too worried about the insignificance of the results. In fact, one way of interpreting the insignificant results is that changes in the urbanization rate are a very noisy proxy for economic growth and insignificant estimates may therefore not necessarily mean that there is no significant relationship between economic growth and cultural development.

An interesting issue that we can shed light on with our output per capita data is whether the marginal effect of polarization and fractionalization on economic growth is particularly large in counties with initially high levels of output per capita. Alesina and LaFerrara (2005) argued that the benefits from variety in skills are more likely to be relevant for more advanced societies. Columns (1)-(4) of Table 6 show that consistent with the argument in Alesina and LaFerrara the positive marginal effect of cultural fractionalization on output growth is particularly large in counties with above median output per capita. In these counties a 1 percentage point increase in cultural fractionalization increased output per capita by up to 4.2 percent. Interestingly, for the above median per capita output counties the negative effect of polarization on output growth is also particularly large. Columns (5)-(8) show on the other hand that quantitatively the effect of fractionalization and polarization on output growth is much smaller and statistically insignificant in the sample with below-median output per capita.[14]

Table 7 presents evidence on an intermediate channel through which cultural fractionalization may have a positive effect on output growth: the diversity of skills of the working-age population. To measure the diversity of skills of the working-age population we constructed a fractionalization index of job occupations as $OCC_{cs} = 1 - \sum_{o=1}^{O} O_{ocs}^2$, where O_{ocs} is the fraction of workers living in US county c in state s that work in occupation o. Economically, this fractionalization index of job occupations captures the probability that two randomly selected workers occupy a different job. Columns (1)-(4) of Table 7 show that once instrumental variables techniques are used there is a significant positive average effect of increases in cultural fractionalization on the fractionalization of job occupations. To reinforce this result we also built another fractionalization measure which captures the probability that two randomly selected individuals work in a different industry. Analog to the fractionalization index of job oc-

[14] In Appendix Table 3 we show that we obtain similar results when splitting countries into above and below median manufacturing GDP per capita. Appendix Table 3 hence confirms that the positive effect of diversity on output works also through manufacturing.

cupations we define the industry diversity index as $IND_{cs} = 1 - \sum_{m=1}^{M} I_{mcs}^2$, where I_{mcs} is the fraction of workers living in US county c in state s that work in industry m. The estimates of columns (5)-(8) indicate that changes in cultural fractionalization are positively associated with changes in the industry diversity index. Hence, both results are consistent with the argument in the cultural economics literature that cultural diversity may have a positive effect on output per capita because it increases the variety of skills, and therefore, the variety of goods and services that are produced in an economy (e.g. Alesina and LaFerrera, 2005).

We also explore the product diversity channel more directly by constructing a fractionalization measure of agricultural diversity as $AGRDIV_{cs} = 1 - \sum_{a=1}^{A} G_{acs}^2$, where G_{acs} is the county product share of good a produced in county c of state s.[15] This fractionalization index captures a county's agricultural production structure and is maximized if a county is completely diversified. Columns (9)-(12) show that increases in the cultural fractionalization of counties were positively associated with a greater variety of agricultural products once instrumental variable techniques are used. These results are yet further evidence that cultural fractionalization fosters economic growth through the product diversity channel.

The conflict literature has argued that increases in cultural polarization may be associated with increases in conflict potential.[16] To explore this channel empirically with our data we present in columns (1)-(4) of Table 8 estimates of the effect that changes in cultural fractionalization and cultural polarization have on changes in the share of people that work in jobs related to crime delicts (policemen, firemen, guards, lawyers, undertakers, sheriffs and marshals). Due to demand effects, it is plausible that the fraction of people which are employed in jobs related to crime delicts is larger, the stronger the underlying conflict potential in the county. We also explore the conflict channel more directly by presenting in columns (5)-(8) estimates of the effect that changes in cultural fractionalization and cultural polarization have on changes in the number of riots. Consistent with the recent conflict literature that has emphasized cultural polarization rather than cultural fractionalization to increase the intensity of conflict, we find that more cultural polarization increased the share of people that worked in jobs related to crime delicts, and that more cultural polarization also increased the number of riots in US counties. Importantly, our estimates are statistically significant when we use instrumental variables estimation to correct for potential endogeneity biases.

To document the significant mean-reversion of the fractionalization and polarization index, Panel A and Panel B of Table 9 show the first-stage estimates where the 1870-1920 change in the fractionalization and polarization index is regressed on the initial 1870 value. The estimated coefficient on the initial polarization index in Panel A is negative and statistically significant at the 1 percent level. Hence, counties that had relatively high levels of polarization

[15] See also the Data Appendix A3 for further details.

[16] See e.g. Esteban and Ray (1999, 2010) or Montalvo and Reynal-Querol (2005b).

in 1870 experienced smaller changes over the 1870-1920 period in polarization than counties that started out with relatively low levels of polarization. Panel B shows that there was also significant reversion to the mean in the fractionalization index. Even though we do not offer here an economic explanation for the mean-reversion in the polarization and fractionalization index, we note that statistically the mean-reverting nature of the polarization and fractionalization index follows from the boundedness of these indices on the unit interval.

An important implication of the significant mean-reversion in the polarization and fractionalization index is that the reduced form effect of initial polarization and initial fractionalization on economic growth has the opposite sign than a two-stage least squares estimate, where the change in polarization and fractionalization is instrumented by the initial value. We show this in Panel C, where the growth rate of output per capita is regressed on the initial values of polarization and fractionalization. The result is that the reduced form coefficients have exactly the opposite sign than what the two-stage least squares estimates revealed in the previous tables. The reason for the discrepancy is that the two-stage least squares estimates, via the first-stage, fully take into account the mean-reversion of the polarization and fractionalization index while the reduced form regression misses this information. For empirical research purposes, we therefore note that using initial measures of fractionalization and polarization in a reduced form regression can produce confounding results.

6 Further Issues

6.1 Group Aggregation

A further issue in our empirical analysis is the level of aggregation of the cultural groups. In contrast to cross-country studies that use measures of ethnic fractionalization, our analysis of the effects of cultural diversity is based on well-defined groups: immigrants' country of origin. By using immigrants' country of origin to construct the cultural diversity measures, our empirical analysis codes cultural groups at a fairly disaggregated level. To check on the importance of disaggregation, we report in Table 10 estimates when we do not distinguish immigrants by country of origin; i.e. when we pool all immigrants into one group so that the fractionalization and the polarization index is based on three groups only: native whites, Afro-Americans, and foreign-born. One disadvantage in terms of identification is that pooling immigrants into one group impedes exploiting network effects to compute the supply-push component of immigrant inflows. This is because network efforts are specific to immigrants' country of origin. We report in Table 10 therefore instrumental variables estimates that use the initial fractionalization and polarization measures to predict the 1870-1920 change (since we cannot use the supply push component as an instrumental variable when we pool immigrants into one group). We also report least-squares estimates.

The main result is that the sign of the estimated coefficients is similar to the

sign of the coefficients that we report in the other tables where we distinguish immigrants by country of origin. Statistically and quantitatively the effects of fractionalization and polarization are however much weaker when we pool immigrants into one group. This should not be surprising if an expectedly large share of the variation in product variety and conflict tensions comes from differences between immigrants, as indeed the anecdotal evidence that we provide in Appendix A1 suggests. When we pool immigrants into one group, these differences are ignored and hence the effects of fractionalization and polarization are quantitatively and statistically weaker.

6.2 Group Distances

In this section we address the issue that our polarization and fractionalization indices are based on the binary criteria of "belonging" or "not belonging" to a particular cultural group. Because using such a discrete metric (belonging/not belonging) implies that the distance across cultural groups is the same, it is the size of each group that matters for our indices. For our research purposes, we believe that it is reasonable to rely on this binary criteria of "belonging" or "not belonging to", because we distinguish cultural groups by the country of origin. We would like to point out that our discrete polarization measure is related to Esteban and Ray's (1994) class of polarization measures for variables that have a continuous dimension. The measurability of inter-group distances on the real line complicates the application of Esteban and Ray's (1994) polarization measure, since it is not obvious how to map accurately inter-group distances on to the real line.

Fearon (2003) and more recently Desmet et al. (2009a) and Esteban et al. (2010) try to tackle this measurement issue by proxying inter-group distances with linguistic distances. To see what we would obtain if we were to follow their approach, we report in Table 11 the results for constructing the fractionalization and polarization index based on linguistic distances.[17] Columns (1)-(4) show that polarization has a significant negative effect on output, and that fractionalization has a significant positive effect. Hence, we continue to obtain our main result that, increases in cultural polarization have a negative output effect while increases in fractionalization have a positive output effect. Importantly, columns (5)-(8) document that the polarization index continues to have a negative effect when we add a quadratic term of the fractionalization index to the right-hand side of the estimating equation.

7 Conclusion

In this paper we investigated the impact that cultural diversity had on economic growth by exploiting the wave of mass immigration to the US during the 1870-1920 period as a plausibly exogenous source of variation in cultural

[17] See the note to Table 11 for details on how exactly these indices are constructed.

diversity. As immigrants flocked from European countries to the US the cultural composition of the US population changed significantly. We showed that increases in the cultural fractionalization of US counties significantly increased output per capita, while increases in the cultural polarization had a significant negative output effect. Our finding that increases in cultural fractionalization significantly increased output per capita is consistent with the argument in the cultural economics literature that greater diversity of the population increases output per capita due to greater variety of workers' skills and a greater variety of goods and services. On the other hand, our finding that increases in cultural polarization significantly decreased output per capita is consistent with the conflict literature that has emphasized the negative, socially destabilizing effect of cultural polarization. An important implication of the findings in our paper is that it is crucial to distinguish between fractionalization and polarization when examining empirically the effects that cultural diversity has on economic growth.

References

Alesina, Alberto, Reza Baqir, and William Easterly, "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114 (1999), 1243-84.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg, "Fractionalization," *Journal of Economic Growth*, 8 (2003), 155-194.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 43 (2005), 762-800.

Atack, Jeremy, Michael R. Haines, and Robert A. Margo, "Railroads and the Rise of the Factory: Evidence for the United States, 1850-70," NBER Working Paper No. 14410, (2008).

Barde, Robert, Susan B. Carter, and Richard Sutch, "International Migration," *Historical Statistics of the United States, Millennial Edition On Line*, Susan B. Carter, Scott Sigmund Gartner, Michael R. Heines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, eds. (Cambridge University Press, 2006).

Bartel, Ann P., "Where Do the New U.S Immigrants Live?" *Journal of Labor Economics*, 7 (1989), 371-391.

Boustan, Leah Platt, "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration," *Quarterly Journal of Economics*, 125 (2010), 417-443.

Card, David, and John DiNardo, "Do Immigrant Inflows Lead to Native Outflows?" *American Economic Review: Paper and Proceedings*, 90 (2000), 360-367.

Card, David, "Immigration Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19 (2001), 22-64.

Card, David, "Immigration and Inequality," *American Economic Review: Paper and Proceedings*, 99 (2009), 1-21.

Cohn, Raymond L., "Immigration to the United States". *EH.Net Encyclopedia*, (2010). URL: http://eh.net/encyclopedia/article/cohn.immigration.us.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Shlomo Weber,"Linguistic Diversity and Redistribution, " *Journal of the European Economic Association*, 7 (2009a), 1291-1318.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Romain Wacziarg,"The Political Economy of Ethnolinguistic Cleavages," NBER Working Paper No. 15360, (2009b).

Easterly, William, and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Division," *Quarterly Journal of Economics*, 112 (1997), 1203-50.

Esteban, Joan, Laura Mayoral, and Debraj Ray, *Ethnicity and Conflict: An Empirical Study.* (Mimeo, NYU, 2010).

Esteban, Joan, and Debraj Ray, "On the Measurement of Polarization," *Econometrica* 62 (1994), 819-851.

Esteban, Joan, and Debraj Ray, "Conflict and Distribution," *Journal of Economic Theory*, 87 (1999), 379-415.

Esteban, Joan, and Debraj Ray, "Linking Conflict to Inequality and Polarization," *American Economic Review*, (2010), forthcoming.

Fearon, James D., "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, 8 (2003), 195-222.

Galor, Oded, Omer Moav, and Dietrich Vollrath, "Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence," *Review of Economic Studies*, 76 (2009), 143-179.

Goldin, Claudia, "The Political Economy of Immigration Restriction in the United States, 1890 to 1921," *NBER Chapters, in: The Regulated Economy: An Historical Approach to Political Economy, pages 223-258*, Claudia Goldin and Gary Libecap, eds. (University of Chicago Press, 1994).

Greenberg, Joseph H., "The Measurement of Linguistic Diversity," *Language*, 32 (1956), 109-115.

Haines, Michael R., and Inter-university Consortium for Political and Social Research, "Historical, Demographic, Economic and Social Data: The United States, 1790-2002." [Computer File]. ICPSR 02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributer], 2010-05-21.

Hatton, Timothy J., and Jeffrey G. Williamson, *The Age of Mass Migration: Causes and Economic Impact* (New York, NY: Oxford University Press, 1998).

Horowitz, Donald L. *Ethnic Groups in Conflict.* (Berkeley: University of California Press, 1985).

Hutchinson, Edward P. *Immigrants and Their Children, 1850-1950.* (New York: John Wiley & Sons, 1956).

Kim, Sukkoo, "Immigration, Industrial Revolution and Urban Growth in the United States, 1820-1920: Factor Endowments, Technology and Geography," NBER Working Paper No. 12900, (2007).

Lane, Philip R., and Aaron Tornell, "Power, Growth, and the Voracity Effect," *Journal of Economic Growth* 1 (1996), 213-241.

Mauro, Paolo, "Corruption and Growth," *Quarterly Journal of Economics*, 60 (1995), 681-712.

Montalvo, Jose G., and Marta Reynal-Querol, "Religious Polarization, and Economic Development," *Economics Letters*, 80 (2003): 201-210.

Montalvo, Jose G., and Marta Reynal-Querol, "Ethnic Diversity and Economic Development," *Journal of Development Economics*, 76 (2005a), 293-323.

Montalvo, Jose G., and Marta Reynal-Querol, "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 95 (2005b), 796-816.

Montalvo, Jose G., and Marta Reynal-Querol, "Discrete Polarization with an Application to the Determinants of Genocides," *Economic Journal*, 118 (2008), 1835-65.

Munshi, Kaivan, "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," Quarterly Journal of Economics 118 (2003), 549-599.

Ottaviano, Gianmarco and Giovanni Peri, "Cities and Cultures," *Journal of Urban Economics*, 58 (2005), 304-337.

Ottaviano, Gianmarco and Giovanni Peri, "The Economic Value of Cultural Diversity: Evidence from US Cities," *Journal of Economic Geography*, 6 (2006), 9-44.

Peri, Giovanni, "The Effect of Immigration on Productivity: Evidence from U.S. States," *Review of Economics and Statistics*, (2010), forthcoming.

Reynal-Querol, Marta, "Ethnicity, Political Systems, and Civil Wars," *Journal of Conflict Resolution*, 46 (2002), 29-54.

Roth Randolph, Douglas L. Eckberg, Cornelia Hughes Dayton, Kenneth Wheeler, James Watkinson, Robb Haberman, and James M. Denham, "The Historical Violence Database: A Collaborative Research Project on the History of Violent Crime and Violent Death," *Historical Methods*, 41 (2008), 81-98.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

Saiz, Albert, "Room in the Kitchen for the Melting Pot: Immigration and Rental Prices," *Review of Economics and Statistics*, 85 (2003), 502-521.

Sparber, Chad, "Racial Diversity and Aggregate Productivity in US Industries: 1980-2000," *Southern Economic Journal*, 75 (2009), 829-856.

Sparber, Chad, "Racial Diversity and Macroeconomic Productivity across US States and Cities," *Regional Studies*, 44 (2010), 71-85.

Stock, James H., and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression." *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Donald Andrews and James H. Stock, eds. (New York, NY: Cambridge University Press, 2005).

Tornell, Aaron, and Philip R. Lane, "The Voracity Effect," *American Economic Review*, 89 (1999), 22-46.

Appendix

A1. Anecdotal Evidence

Diversity of the Workforce

During the 1870-1920 period, the occupational distribution reveals a striking pattern of occupational clustering by nationalities (Hutchinson, 1956).[18] As Kim (2007, pp. 17-18) pointed out:

> The Germans specialized in many food related industries as brewers, distillers, butchers and confectioners ... The Irish were highly specialized in gas works and other heavy industries ... The English and Welsh were concentrated in textiles, iron and steel ... The Scandinavians were highly specialized in a few occupations: sail and awning makers and those related to lumber industry.

The occupational clustering by nationalities provides evidence in support of the product variety channel of cultural diversity emphasized by Alesina and La Ferrara (2005), and is in line with our empirical findings of Table 7 (Section 5, page 11).

Cultural Polarization

Social tensions among European immigrants were not rare in the US during the 19th century. Antagonism between Europeans that was due to the rise of nationalism in 19th century Europe, was at least in part imported to the US by the large inflow of European immigrants. The following examples illustrate some of the cultural conflicts among European immigrants in the US during the 1870-1920 period.

In 1868 the *The New York Times* reported a riot between German and Irish immigrants that ended up with thirty men wounded and sixty arrested by the police at Ward's Island, New York. The dispute between the Irish and Germans started, when "the contestants used vile epithets toward each other's nationality".[19] Another riot took place in Scranton, Pennsylvania, 1871 among Welsh, Irish and German coal miners. Violence occurred among the strike members when German and Irish miners were attacked by Welsh strikers.[20] In Pittsburg, Pennsylvania, 1886 a conflict between Irish and Italian laborers – residing in the same neighborhood – arose, "in which two of the participants received fatal injuries".[21] A severe fight among Swedish, Polish and Hungarian immigrants in Denver, Colorado, 1887 resulted in one man being shot and several others

[18] See Kim (2007) for a detailed descriptive statistic on occupational clustering of foreign-born worker by nationalities for the years 1870 and 1890.

[19] *The New York Times*, published: March 6, 1868 as *"Riot on Ward's Island"*.

[20] *The New York Times*, published: May 11, 1871 as *"The Coal Riot"*.

[21] *The New York Times*, published: September 20, 1886 as *"Fatal War Among Races"*.

seriously wounded.[22] In 1915, a clash between Italians and Austrians occurred at the Federal Pressed Steel Company in Milwaukee, Wisconsin. Seven men were hurt after the Austrian steel workers left because the workers discovered that they were hired by the company to keep up with orders for shrapnel which were sold to Russia.[23]

A2. Simulation

To illustrate and provide some intuition for the relationship between the fractionalization and polarization index, we simulated data on the population share of n groups by drawing n times from a uniform [0,1] distribution and dividing each draw by the total size of the realizations drawn. We repeated this 1000 times, constructing at each replication the fractionalization and polarization index. Appendix Figure 1 shows the results for $n = 3, 4, 5, 6, 7$. The key message is that, when the number of groups is small the polarization and fractionalization index are positively correlated and as the number of groups increase the correlation turns negative.[24]

A3. Data Appendix

Our main data source is the Inter-University Consortium for Political and Social Research (ICPSR) 2896 data file. The ICPSR 2896 data file contains detailed decennial county and state level data on demographic, economic, and social variables which were collected by the U.S. Bureau of Census for the period 1790-2000.[25] One key advantage of the ICPSR data set is that it enables us to exploit the underlying cultural heterogeneity in the United States at the county level. In particular, the ICPSR data set comprises – from 1870 onwards – detailed information about the country of origin of foreign-born, white US-born citizens and Afro-Americans which is necessary to calculate the cultural diversity indices described in Section 4.

As a further database, we use the Integrated Public Use Microdata Series (IPUMS-USA) which consists of more than fifty high-precision samples of the American population drawn from fifteen federal censuses and from the American Community Surveys of 2000-2008. The IPUMS is composed of micro-data, where each record is a person, with all characteristics numerically coded. In most samples persons are organized into households, making it possible to study the characteristics of people in the context of their families or other co-residents. These samples, which draw on every surviving census from 1850-2000, and the

[22] *The New York Times*, published: April 12, 1887 as *"A Race Riot in Denver"*.

[23] *The New York Times*, published: August 16, 1915 as *"Seven Hurt in Race Riot"*.

[24] We did not show the results for $n = 2$ because in this case the fractionalization and polarization index are exactly the same (see Montalvo and Reynal-Querol (2005a,b)).

[25] More information about the data set (i.e. scope of study, data collection and data source) can be found at http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/02896.

2000-2008 ACS samples, collectively constitute a rich source of quantitative information on long-term changes in the American population.[26] The IPUMS gives us the possibility to exploit individual level data and construct aggregate data at the county level whenever these variables are missing in the ICSPR data set, but available at the IPUMS. In particular, we used the IPUMS database to construct our fractionalization index of job occupations, the industry diversity index and the share of people that work in jobs related to crime delicts (see Section 5). The IPUMS benchmark occupation classification variable *occ1950* which assigns an occupation code to the individual's reported occupation is used to construct the fractionalization index of job occupations and the share of people that work in jobs related to crime delicts. To build our industry diversity index, we exploit the IPUMS benchmark industry classification variable *ind1950*, which contains coded information about the industry in which an individual performed an occupation.[27]

Our source for the number of riots in 1870 and 1920 is the historical violence database (HVD).[28] The dataset reports detailed information on riots (year, kills, location and a further description of each riot) in the United States from colonial times to the present. Although the database reports the location (usually either the city, town or county) of the riots, we had to compile for each location the corresponding (ICSPR) county code (if the information on the location was precise enough) to make the data source for our analysis compatible. The following table provides a detailed description of the variables used in our empirical analysis (if not further specified, variables are selected from the ICSPR 2896 data file):

[26]The IPUMS is a public project and data are freely available. For more information see: http://usa.ipums.org/usa/index.shtml.

[27]The occupation classification code of 1950 is a reference for all Census occupation data available at the IPUMS. Census years with different occupation coding schemes (as the 1870 and 1920 Census) are converted by the IPUMS into the occupation classification of 1950 (i.e. the variable *occ1950*) to make occupations over time comparable. The same applies to the IPUMS industrial classification variable *ind1950*. More information about the creation of both variables *ind1950* and *occ1950* is provided by the IPUMS at http://usa.ipums.org/usa/index.shtml.

[28]The HVD is a public project and data are freely available. For more information see: http://cjrc.osu.edu/researchprojects/hvd/usa/riots/ and Roth et al. (2008).

DEPENDENT VARIABLES

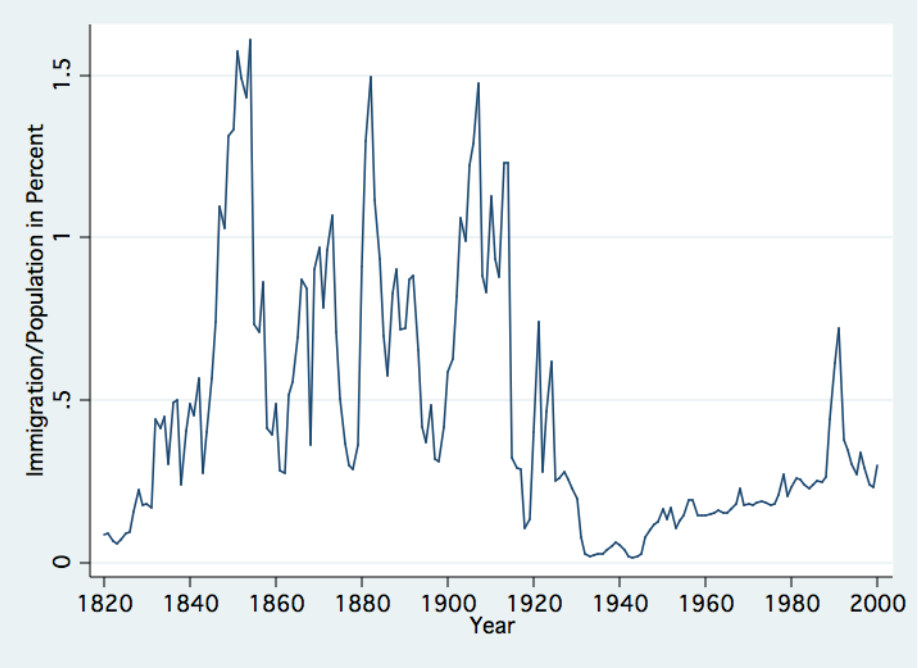
VARIABLE	YEARS	DESCRIPTION
Output growth	1870 - 1920	Total output in per capita terms is formed as the sum of manufacturing output and agricultural output. The growth variable is calculated as the change in logarithmic units.
Manufacturing Output	1870 - 1920	Manufacturing output is in per capita terms. For 1910, there are no manufacturing Census data available at the county level.
Agricultural Output	1870 - 1920	Agricultural output is in per capita terms. For 1870 we use the variable "farmout", which contains the estimated value of farm products. For 1920 we add the values of crops, cereals, other grains and seeds, hay and forage, vegetables, fruits and all other crops reported in the ICSPR 2896 data file.
Urban growth	1870 - 1920	Change in the population share living in urban counties. The Census declared a county population as urban, if at least 2500 inhabitants lived in urban places.
Change in Occupational Diversity Index	1870 - 1920	We take the occupation variable *occ1950* from the IPUMS to construct the occupational diversity index. See the last paragraph of page 11 in the paper for further details.
Change in Industry Diversity Index	1870 - 1920	We take the industry classification variable *ind1950* from the IPUMS to construct the industry diversity index. See the last paragraph of page 11 in the paper for further details.

DEPENDENT VARIABLES (CONTINUED)		
Change in Agricultural Diversity Index	1870 - 1920	We use the following agricultural goods for our index: wheat, rye, corn, oats, barley, buckwheat, peas and beans, potato, sweet potato, emmer and spelt, kafir corn, soybeans and peanuts which are all measured in bushels. These agricultural goods are reported in the ICSPR 2896 data file Part 47 (1870) and Part 86 (1920). See the first paragraph of page 12 in the paper for further details.
Change in Occupations Related to Crime Delicts	1870 - 1920	We consider the following occupations as related to crime delicts: sheriffs, marshals, policemen, firemen, guards, undertaker and lawyers. These occupations are listed in the Census occupational classification system of 1950 and reported by the IPUMS, separately (see variable *occ1950*). We group these occupations together and calculate the corresponding share as occupations related to crime delicts out of the total workforce.
Change in Number of Riots	1870 - 1920	We construct the number of riots in US counties from the HVD.

ADDITIONAL CONTROLS		
VARIABLE	YEARS	DESCRIPTON
Land Concentration	1870	Gini coefficient of farm size distribution, calculated as in Galor et al. (2009, p. 175).
Manufacturing Share	1870	Share of manufacturing output over the sum of manufacturing and agricultural output in 1870.
Population	1870	Total population in US counties in 1870.
Rail Access	1870	Indicator variable that is equal to one if a county has access to a railroad within its borders in 1870, and zero otherwise. See Atack et al. (2008) for more information on the construction of the railroad data set.
Share of Foreign-Born	1870	Share of foreign-born out of the total population in 1870.
Share of Afro-Americans	1870	Share of Afro-Americans (declared by the historical US Census as negro population) out of the total population in 1870.

Tables and Figures

Figure 1: Annual Immigration as a Fraction of the US Population, 1820 - 2000



Source: Historical Statistics of the United States (Barde et al., 2006), Statistical Abstract of the US, eh.net database and Kim (Figure 3, 2007).

Figure 2: Fractionalization vs. Polarization

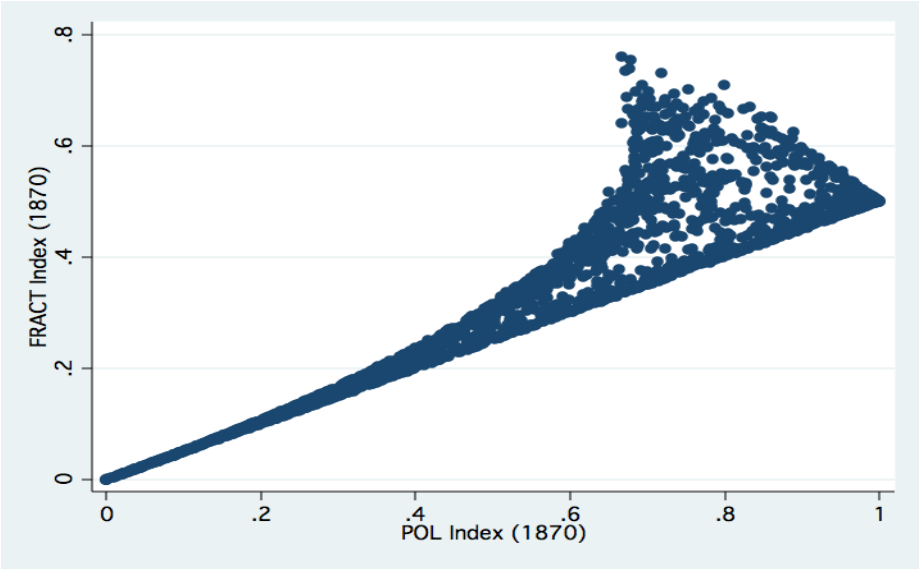


Figure 2A: Year 1870

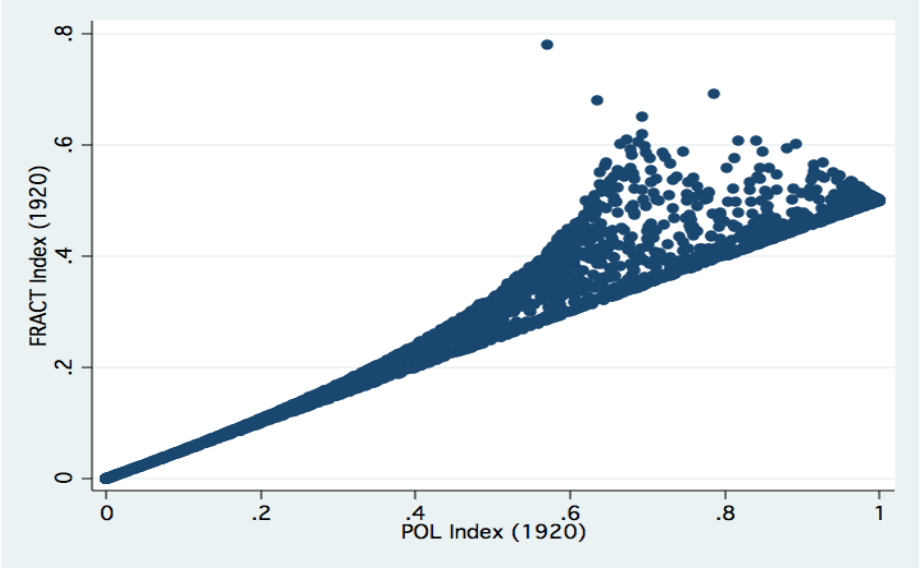


Figure 2B: Year 1920

Table I
Foreign-Born Descriptive Statistics (1870 - 1920)

VARIABLES	Mean	Min	Max	STANDARD DEVIATION		
				Overall	Between	Within
Foreign-Born (FB), (overall)	0.10	0.00	1.00	0.12	0.11	0.04
FB in the West	0.19	0.00	0.95	0.11	0.09	0.06
FB in the Midwest	0.15	0.00	1.00	0.13	0.12	0.06
FB in the Northeast	0.15	0.00	0.54	0.10	0.09	0.04
FB in the South	0.02	0.00	0.45	0.05	0.05	0.02
Afro-Americans, (overall)	0.13	0.00	1.00	0.21	0.20	0.03
White US-Born, (overall)	0.76	0.00	1.00	0.19	0.19	0.05

Table II
Fractionalization & Polarization Index: Descriptive Statistics (1870 - 1920)

FRACTIONALIZATION							
				MEAN (MACROREGIONS)			
YEAR	Mean	25^{th} Percentile	75^{th} Percentile	West	South	Midwest	Northeast
1870	0.3302	0.1793	0.4818	0.4161	0.3491	0.3055	0.2539
1880	0.3138	0.1700	0.4635	0.3899	0.3313	0.2919	0.2394
1890	0.3153	0.1674	0.4661	0.3719	0.3202	0.3044	0.2672
1900	0.2901	0.1419	0.4393	0.3172	0.3105	0.2611	0.2682
1910	0.2877	0.1342	0.4339	0.3214	0.3022	0.2480	0.3303
1920	0.2572	0.1120	0.4025	0.2540	0.2940	0.1973	0.3048
Overall	*0.2970*	*0.1479*	*0.4489*	*0.3323*	*0.3163*	*0.2663*	*0.2774*

POLARIZATION							
				MEAN (MACROREGIONS)			
YEAR	Mean	25^{th} Percentile	75^{th} Percentile	West	South	Midwest	Northeast
1870	0.5633	0.3319	0.7967	0.5660	0.6668	0.4680	0.4188
1880	0.5442	0.3165	0.7682	0.5590	0.6367	0.4607	0.3989
1890	0.5344	0.3098	0.7404	0.5268	0.6124	0.4677	0.4247
1900	0.5065	0.2657	0.7205	0.4766	0.5975	0.4214	0.4209
1910	0.5003	0.2540	0.7074	0.4772	0.5821	0.4015	0.4975
1920	0.4595	0.2129	0.6676	0.4052	0.5681	0.3332	0.4692
Overall	*0.5151*	*0.2756*	*0.7341*	*0.4885*	*0.6076*	*0.4232*	*0.4384*

Table III
The Impact of Polarization and Fractionalization on Output Growth

	Output p.c. (log-change) 1870 - 1920						
	(1) LS	(2) LS	(3) LS	(4) LS	(5) IV (A)	(6) IV (B)	(7) IV (C)
Change POL 1870-1920	-0.237*		-1.081***	-1.139***	-1.708**	-2.525***	-1.968***
	(0.133)		(0.408)	(0.273)	(0.676)	(0.611)	(0.451)
Change FRAC 1870-1920		-0.132	1.359*	1.578***	1.911**	2.278***	2.017***
		(0.247)	(0.720)	(0.446)	(0.921)	(0.673)	(0.599)
Output p.c. 1870				-0.708***	-0.706***	-0.703***	-0.705***
				(0.0415)	(0.0409)	(0.0405)	(0.0406)
Land Concentration 1870				-0.590***	-0.545***	-0.474***	-0.522***
				(0.143)	(0.146)	(0.149)	(0.146)
Manufacturing Share 1870				-0.177**	-0.169**	-0.161**	-0.167**
				(0.0800)	(0.0778)	(0.0794)	(0.0776)
Population 1870				0.0489***	0.0452**	0.0395**	0.0433**
				(0.0185)	(0.0189)	(0.0193)	(0.0191)
Rail Access 1870				0.0749***	0.0804***	0.0901***	0.0837***
				(0.0277)	(0.0285)	(0.0289)	(0.0283)
Share Foreign-Born 1870				0.0621***	0.0505***	0.0302	0.0437***
				(0.0134)	(0.0179)	(0.0190)	(0.0169)
Share Afro-Americans 1870				-0.00739	-0.00848	-0.0105	-0.00915
				(0.0104)	(0.0104)	(0.0105)	(0.0104)
Observations	2068	2068	2068	2068	2068	2068	2068
R^2	0.264	0.262	0.269	0.543	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	-	-	-	11.01	42.81	14.17
Hansen-J statistic (p-val.)	-	-	-	-	0.82	n.a.	0.79
Anderson-Rubin Wald-Test (p-val.)	-	-	-	-	0.00	0.00	0.00
Endogeneity Test Statistic (p-val.)	-	-	-	-	0.04	0.01	0.00
State FE	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In columns (1)-(4) the method of estimation is least squares. In columns (5)-(7) the method of estimation is two-stage least squares. In column (5) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In column (6) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In column (7), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)).

Table IV
The Impact of Polarization and Fractionalization on Output Growth: Different Time Periods

PANEL A: Output p.c. (log-change) 1880 - 1920				
	LS	IV (A)	IV (B)	IV (C)
Change POL 1880-1920	-1.055***	-0.788	-2.172***	-1.726***
	(0.303)	(0.907)	(0.694)	(0.585)
Change FRAC 1880-1920	1.582***	1.439	2.766***	2.496***
	(0.499)	(1.205)	(0.857)	(0.807)
Observations	2331	2331	2331	2331
R^2	0.578	-	-	-
PANEL B: Output p.c. (log-change) 1890 - 1920				
	LS	IV (A)	IV (B)	IV (C)
Change POL 1890-1920	-1.746***	-1.373**	-2.553***	-2.182***
	(0.289)	(0.691)	(0.657)	(0.492)
Change FRAC 1890-1920	2.251***	1.886**	2.734***	2.619***
	(0.474)	(0.797)	(0.665)	(0.668)
Observations	2438	2438	2438	2438
R^2	0.545	-	-	-
PANEL C: Output p.c. (log-change) 1900 - 1920				
	LS	IV (A)	IV (B)	IV (C)
Change POL 1900-1920	-1.197**	-1.889	-4.966***	-4.070***
	(0.524)	(1.306)	(1.321)	(0.862)
Change FRAC 1900-1920	1.779**	1.567	5.196***	4.070***
	(0.840)	(1.601)	(1.310)	(1.169)
Observations	2614	2614	2614	2614
R^2	0.371	-	-	-

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In column (1) the method of estimation is least squares. In columns (2)-(4) the method of estimation is two-stage least squares. In column (2) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In column (3) the instrumental variables are the initial polarization and fractionalization index (IV (B)). In column (4), the instrumental variables are the supply-push component of immigrant inflows and the initial polarization and fractionalization index (IV (C)). Initial control variables are output per capita, land concentration, the manufacturing share, population size, the share of foreign-born and the share of Afro-Americans (estimates not reported in the table; since we only have information on counties' rail access in 1870, rail access is not included as additional control). For Panel A-C, the first-stage (Kleibergen-Paap) F-Statistic is between 11.97 - 20.08 for column (2), between 24.29 - 42.91 for col. (3) and between 11.78 - 22.00 for column (4). In every specification, the p-value of the Anderson-Rubin Wald-Test is smaller than 0.01. For column (2) and (4) in Panel A, the p-value of the Hansen-J statistic is 0.17/0.38, in Panel B the p-value is 0.38/0.23 and in Panel C the p-value is 0.18/0.14, respectively.

Table V
The Impact of Polarization and Fractionalization on Urban Growth

	Change in the Urban Share 1870 - 1920						
	(1) LS	(2) LS	(3) LS	(4) LS	(5) IV (A)	(6) IV (B)	(7) IV (C)
Change POL 1870-1920	0.208***		-0.0106	-0.0598	-0.105	-0.0305	-0.00733
	(0.0306)		(0.0873)	(0.0879)	(0.226)	(0.178)	(0.139)
Change FRAC 1870-1920		0.337***	0.352**	0.620***	0.391	0.246	0.273
		(0.0521)	(0.146)	(0.151)	(0.257)	(0.192)	(0.174)
Urban Share 1870				-0.330***	-0.320***	-0.322***	-0.325***
				(0.0452)	(0.0435)	(0.0446)	(0.0434)
Land Concentration 1870				-0.0956*	-0.0766	-0.0759	-0.0801
				(0.0491)	(0.0508)	(0.0492)	(0.0495)
Output p.c. 1870				-0.0200**	-0.0170*	-0.0161*	-0.0166*
				(0.00955)	(0.00965)	(0.00960)	(0.00956)
Manufacturing Share 1870				0.158***	0.147***	0.144***	0.145***
				(0.0295)	(0.0304)	(0.0305)	(0.0302)
Population 1870				0.0384***	0.0366***	0.0369***	0.0373***
				(0.00880)	(0.00816)	(0.00849)	(0.00823)
Rail Access 1870				0.0212**	0.0259***	0.0268***	0.0258***
				(0.00946)	(0.00967)	(0.00961)	(0.00956)
Share Foreign-Born 1870				0.0367***	0.0271***	0.0253***	0.0272***
				(0.00444)	(0.00589)	(0.00574)	(0.00534)
Share Afro-Americans 1870				0.00243	0.00132	0.00111	0.00133
				(0.00385)	(0.00391)	(0.00396)	(0.00390)
Observations	2068	2068	2068	2068	2068	2068	2068
R^2	0.203	0.207	0.207	0.308	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	-	-	-	11.85	40.86	14.77
Hansen-J statistic (p-val.)	-	-	-	-	0.29	n.a.	0.37
Anderson-Rubin Wald-Test (p-val.)	-	-	-	-	0.00	0.20	0.00
Endogeneity Test Statistic (p-val.)	-	-	-	-	0.04	0.00	0.00
State FE	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In columns (1)-(4) the method of estimation is least squares. In columns (5)-(7) the method of estimation is two-stage least squares. In column (5) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In column (6) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In column (7), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)).

Table VI
The Impact of Polarization and Fractionalization on Output Growth: Below vs. Above Median GDP per capita

	Output p.c. (log-change) 1870-1920: Columns (1)-(4) Above; Columns (5)-(8) Below Median GDP p.c.							
	(1) LS	(2) IV (A)	(3) IV (B)	(4) IV (C)	(5) LS	(6) IV (A)	(7) IV (B)	(8) IV (C)
Change POL 1870-1920	-1.736***	-2.661***	-3.074***	-1.799***	0.266	0.127	-1.594	-1.295*
	(0.341)	(0.920)	(0.618)	(0.434)	(0.424)	(0.565)	(1.183)	(0.702)
Change FRAC 1870-1920	2.806***	4.231***	3.088***	2.701***	-1.148	-0.893	1.247	0.730
	(0.501)	(1.447)	(0.768)	(0.733)	(0.739)	(0.953)	(1.254)	(1.063)
Output p.c. 1870	-0.577***	-0.572***	-0.531***	-0.572***	-0.704***	-0.706***	-0.731***	-0.727***
	(0.0639)	(0.0609)	(0.0625)	(0.0612)	(0.0755)	(0.0736)	(0.0788)	(0.0770)
Land Concentration 1870	-0.195	-0.228	-0.185	-0.190	-0.818***	-0.815***	-0.676***	-0.686***
	(0.219)	(0.227)	(0.214)	(0.215)	(0.194)	(0.212)	(0.218)	(0.196)
Manufacturing Share 1870	-0.323***	-0.274***	-0.287***	-0.323***	-0.0574	-0.0536	-0.0377	-0.0449
	(0.0948)	(0.0916)	(0.0964)	(0.0919)	(0.154)	(0.143)	(0.145)	(0.141)
Population 1870	0.0614***	0.0579***	0.0516**	0.0606***	0.00332	0.00319	0.000231	0.000550
	(0.0221)	(0.0221)	(0.0229)	(0.0220)	(0.0297)	(0.0288)	(0.0301)	(0.0299)
Rail Access 1870	0.125***	0.118***	0.145***	0.128***	0.0442	0.0433	0.0493	0.0506
	(0.0418)	(0.0415)	(0.0420)	(0.0405)	(0.0351)	(0.0361)	(0.0383)	(0.0363)
Share Foreign-Born 1870	0.108***	0.125***	0.0487	0.0993***	0.00786	0.0101	0.0118	0.00771
	(0.0229)	(0.0346)	(0.0334)	(0.0317)	(0.0165)	(0.0219)	(0.0224)	(0.0203)
Share Afro-Americans 1870	-0.00893	-0.00612	-0.0127	-0.00963	0.0249	0.0250	0.0211	0.0212
	(0.0128)	(0.0137)	(0.0136)	(0.0131)	(0.0156)	(0.0155)	(0.0164)	(0.0151)
Observations	1046	1046	1046	1046	1022	1022	1022	1022
R^2	0.505	-	-	-	0.535	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	7.40	28.72	10.80	-	14.43	11.74	20.25
Hansen-J statistic (p-val.)	-	0.50	n.a.	0.42	-	0.47	n.a.	0.86
Anderson-Rubin Wald-Test (p-val.)	-	0.00	0.00	0.00	-	0.00	0.38	0.00
Endogeneity Test Statistic (p-val.)	-	0.39	0.00	0.07	-	0.47	0.02	0.00
State FE	yes	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In columns (1) and (5) the method of estimation is least squares. In columns (2)-(4) and (6)-(8) the method of estimation is two-stage least squares. In columns (2) and (6) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In columns (3) and (7) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In columns (4) and (8), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)).

Table VII
The Impact of Polarization and Fractionalization on Occupational, Industry and Agricultural Diversity

	Δ Occupational Diversity Index 1870-1920				Δ Industry Diversity Index 1870-1920				Δ Agricultural Diversity Index 1870-1920			
	(1) LS	(2) IV (A)	(3) IV (B)	(4) IV (C)	(5) LS	(6) IV (A)	(7) IV (B)	(8) IV (C)	(9) LS	(10) IV (A)	(11) IV (B)	(12) IV (C)
Change POL 1870-1920	0.126***	-0.144*	-0.103	-0.0501	0.223***	-0.141	0.111	-0.0249	0.0649	-0.316	-0.213	-0.187
	(0.0374)	(0.0838)	(0.0983)	(0.0606)	(0.0648)	(0.138)	(0.171)	(0.107)	(0.0774)	(0.230)	(0.169)	(0.126)
Change FRAC 1870-1920	0.0105	0.264***	0.194**	0.152**	0.0559	0.277*	0.175	0.205*	0.0617	0.630**	0.449***	0.466***
	(0.0540)	(0.0988)	(0.0804)	(0.0694)	(0.0932)	(0.166)	(0.139)	(0.123)	(0.125)	(0.252)	(0.167)	(0.147)
Observations	2046	2046	2046	2046	2046	2046	2046	2046	2061	2061	2061	2061
R^2	0.775	-	-	-	0.497	-	-	-	0.375	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	10.43	41.30	15.99	-	10.43	41.30	15.99	-	11.06	42.90	14.16
Hansen-J statistic (p-val.)	-	0.34	n.a.	0.28	-	0.33	n.a.	0.30	-	0.32	n.a.	0.42
Anderson-Rubin Wald-Test (p-val.)	-	0.00	0.03	0.00	-	0.00	0.03	0.00	-	0.00	0.01	0.00
Endogeneity Test Statistic (p-val.)	-	0.04	0.01	0.01	-	0.03	0.49	0.03	-	0.20	0.00	0.00
State FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In columns (1),(5) and (9) the method of estimation is least squares. In columns (2)-(4), columns (6)-(8) and columns (10)-(12) the method of estimation is two-stage least squares. In columns (2), (6) and (10) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In columns (3), (7) and (11) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In columns (4), (8) and (12) the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)). The initial control variables (1870) for columns (1)-(12) are output per capita, population size, land concentration, the manufacturing share, rail access, the share of foreign-born and the share of Afro-Americans (estimates not reported in the table).

Table VIII
The Impact of Polarization and Fractionalization on Occupations Related to Crime Delicts and Riots

	Δ Crime Related Jobs 1870-1920				Δ Number of Riots 1870-1920			
	(1) LS	(2) IV (A)	(3) IV (B)	(4) IV (C)	(5) LS	(6) IV (A)	(7) IV (B)	(8) IV (C)
Change POL 1870-1920	0.00647*	0.0117	0.0233***	0.0128**	7.768***	8.767***	11.42	10.35***
	(0.00355)	(0.00948)	(0.00780)	(0.00540)	(2.615)	(2.472)	(9.798)	(2.494)
Change FRAC 1870-1920	-0.00117	-0.0144	-0.0242***	-0.0132*	-10.48**	-15.34***	-17.27	-14.24***
	(0.00563)	(0.0119)	(0.00831)	(0.00745)	(4.350)	(2.999)	(12.36)	(3.121)
Observations	2046	2046	2046	2046	44	44	44	44
R^2	0.606	-	-	-	0.391	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	11.30	42.27	17.23	-	7.66	0.73	26.98
Hansen-J statistic (p-val.)	-	0.10	n.a.	0.13	-	0.43	n.a.	0.30
Anderson-Rubin Wald-Test (p-val.)	-	0.00	0.00	0.00	-	0.00	0.58	0.00
Endogeneity Test Statistic (p-val.)	-	0.60	0.00	0.03	-	0.10	0.86	0.23
State FE	yes	yes	yes	yes	no	no	no	no

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In columns (1) and (5) the method of estimation is least squares. In columns (2)-(4) and columns (6)-(8) the method of estimation is two-stage least squares. In columns (2) and (6), the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In columns (3) and (7), the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In columns (4) and (8), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)). The initial control variables (1870) for columns (1)-(8) are output per capita, population size, land concentration, the manufacturing share, rail access, the share of foreign-born and the share of Afro-Americans (estimates not reported in the table). Note that there are many counties with no observations on riots in the historical violence database (HVD). This lack of observations leads to a substantial drop in the number of observations in columns (5)-(8). State fixed effects are therefore not included. Also see the Data Appendix A3 for further details.

Table IX
Mean-Reversion of the Polarization and Fractionalization Index

PANEL A: Mean-Reversion Polarization Index 1870 - 1920				
	Dependent Variable: ΔPOL 1870 - 1920			
	(1)	(2)	(3)	(4)
Initial POL 1870	-0.184***	-0.147***	-0.247***	-0.182***
	(0.0189)	(0.0290)	(0.0317)	(0.0188)

PANEL B: Mean-Reversion Fractionalization Index 1870 - 1920				
	Dependent Variable: $\Delta FRAC$ 1870 - 1920			
	(1)	(2)	(3)	(4)
Initial FRAC 1870	-0.302***	-0.263***	-0.361***	-0.303***
	(0.0203)	(0.0331)	(0.0307)	(0.0203)

PANEL C: Reduced Form, 1870 - 1920				
	Dependent Variables: see Footnote			
	(1)	(2)	(3)	(4)
Initial POL 1870	0.905***	0.463	1.478***	0.0865
	(0.256)	(0.498)	(0.317)	(0.0736)
Initial FRAC 1870	-1.313***	-0.663	-2.142***	-0.201
	(0.454)	(0.910)	(0.545)	(0.134)
Observations	2068	1022	1046	2068
R^2 (for Panel C)	0.542	0.529	0.496	0.251
State FE	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In Column (1) we report the counterpart to Table III, column (6). Column (2) reports the counterpart to Table VI, column (7). Column (3) reports the counterpart to Table VI, column (3) and in column (4) we report the counterpart to Table V, column (6). Dependent variables for Panel C: columns (1)-(3) output growth; column (4) the change in the urban share. In Panels A-C the initial control variables (1870) for columns (1)-(3) are output per capita, population size, land concentration, the manufacturing share, rail access, the share of foreign-born and the share of Afro-Americans. In addition to these control variables we include in column (4) the initial urban share (estimates not reported in the table).

Table X
The Impact of Polarization and Fractionalization using Three Groups only: 1870 - 1920

	Δ Output p.c.		Δ Occupational Index		Δ Industry Index		Δ Crime Related Jobs		Δ Number of Riots	
	(1) LS	(2) IV	(3) LS	(4) IV	(5) LS	(6) IV	(7) LS	(8) IV	(9) LS	(10) IV
Change POL 1870-1920	-0.0512	-1.840*	0.0529	-0.00816	0.0867	0.0194	0.00179	0.0146	2.017	5.870
	(0.602)	(0.989)	(0.0726)	(0.116)	(0.153)	(0.245)	(0.00609)	(0.0110)	(9.050)	(13.50)
Change FRACT 1870-1920	-0.125	2.618	0.121	0.100	0.272	0.355	0.00621	-0.0256	-0.474	-7.995
	(1.165)	(1.961)	(0.139)	(0.242)	(0.292)	(0.514)	(0.0118)	(0.0239)	(17.36)	(26.82)
Observations	2068	2068	2046	2046	2046	2046	2046	2046	44	44
R^2	0.536	-	0.774	-	0.496	-	0.606	-	0.334	-
First-stage (Kleibergen-Paap) F-Statistic	-	66.91	-	60.39	-	60.39	-	61.92	-	0.87
Hansen-J statistic (p-val.)	-	n.a.	-	n.a.	-	n.a.	-	n.a.	-	n.a.
Anderson-Rubin Wald-Test (p-val.)	-	0.04	-	0.70	-	0.11	-	0.37	-	0.84
Endogeneity Test Statistic (p-val.)	-	0.03	-	0.31	-	0.90	-	0.21	-	0.90
State FE	yes	yes	yes	yes	yes	yes	yes	yes	no	no

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The three groups used for the construction of the polarization and fractionalization index are foreign-born, US-born whites and Afro-Americans. In columns (1), (3), (5), (7) and (9) the method of estimation is least squares. In columns (2), (4), (6), (8) and (10) the method of estimation is two-stage least squares, where the instrumental variables are the initial 1870 polarization and fractionalization index. The initial control variables (1870) for columns (1)-(10) are output per capita, population size, land concentration, the manufacturing share, rail access, the share of foreign-born and the share of Afro-Americans. In addition to these control variables we include: the initial occupational diversity and the initial industry diversity in columns (3)-(6) and in columns (7)-(8) the initial workforce and the initial share of crime related jobs per worker as further control variables (1870) (estimates not reported in the table). Note that there are many counties with no observations on riots in the historical violence database (HVD). This lack of observations leads to a substantial drop in the number of observations in columns (5)-(8). State fixed effects are therefore not included. Also see the Data Appendix A3 for further details.

Table XI
Polarization and Fractionalization Index with Linguistic Distances (LD)

	Output p.c. 1870-1920 (log-change)							
	(1) LS	(2) IV (A)	(3) IV (B)	(4) IV (C)	(5) LS	(6) IV (A)	(7) IV (B)	(8) IV (C)
Change POL_{LD} *1870-1920*	-20.53***	-24.34**	-20.12**	-20.74***	-19.74***	-24.77**	-20.09**	-21.08***
	(4.979)	(10.91)	(8.156)	(6.388)	(6.292)	(10.89)	(8.035)	(6.956)
Change $FRAC_{LD}$ *1870-1920*	8.082***	9.000**	6.887**	7.429***	7.472**	11.35**	7.045*	7.831**
	(2.003)	(4.163)	(3.355)	(2.483)	(3.317)	(5.126)	(4.136)	(3.628)
Change $FRAC^2_{LD}$ *1870-1920*					1.639	-9.792	-0.916	-1.245
					(6.418)	(7.664)	(6.680)	(6.231)
Observations	2068	2068	2068	2068	2068	2068	2068	2068
R^2	0.542	-	-	-	0.542	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	10.81	104.05	315.75	-	8.96	107.57	40.37
Hansen-J statistic (p-val.)	-	0.73	n.a.	0.74	-	0.81	n.a.	0.73
Anderson-Rubin Wald-Test (p-val.)	-	0.00	0.01	0.00	-	0.00	0.02	0.00
Endogeneity Test Statistic (p-val.)	-	0.10	0.17	0.03	-	0.65	0.23	0.09
State FE	yes	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In columns (1) and (5) the method of estimation is least squares. In columns (2)-(4) and (6)-(8) the method of estimation is two-stage least squares. In columns (2) and (6) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In columns (3) and (7) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In columns (4) and (8), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)). Initial control variables (1870) for columns (1)-(8) are output per capita, land concentration, the manufacturing share, population size, rail access, the share of foreign-born and the share of Afro-Americans (estimates not reported in the table). We construct our measure of fractionalization with distances following Greenberg (1956) as: $FRAC_{LD} = \sum_{i=1}^{N} \sum_{j=1}^{N} pi_i pi_j d_{ij}$, where subindices for counties and states are left out for simplicity. The corresponding measure of polarization follows Esteban and Ray (1994, 1999) and is constructed as: $POL_{LD} = \sum_{i=1}^{N} \sum_{j=1}^{N} pi_i^2 pi_j d_{ij}$. Both indices include a measure of inter-group distances d_{ij}. To proxy for inter-group distances, we follow Fearon (2003), Desmet et al. (2009a) and Esteban et al. (2010) and use information on linguistic groups compiled by the Ethnologue project to construct a measure of linguistic distances between any two groups as $d_{ij} = 1 - b_{ij}^{\delta}$.[a] The parameter b_{ij} is the ratio of the number of shared branches between i and j to the maximum number of branches between any two languages and $\delta \in (0, 1]$ represents a sensitivity parameter determining how fast the distance declines as the number of shared branches increases (see Desmet et al., 2009a).[b] The abbreviation LD denotes the use of language distances for the metric d_{ij}. We use the representative language of each country of origin to construct the linguistic distances. Thus, if two groups speak the same representative language we set $b_{ij} = 1$. Following Fearon (2003) we compute the linguistic distances d_{ij} by setting $\delta = 0.5$.[c]

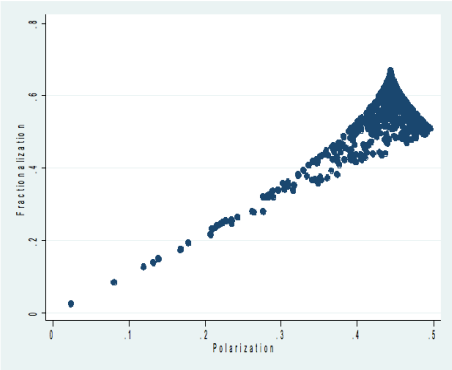
[a] See http://www.ethnologue.com, Desmet et al. (2009a) and Esteban et al. (2010) for further information on the Ethnologue project.

[b] We used the language trees reported by Ethnologue to construct the parameter b_{ij}. See also Desmet et al. (2009b) for a detailed discussion on the construction of such a language tree.

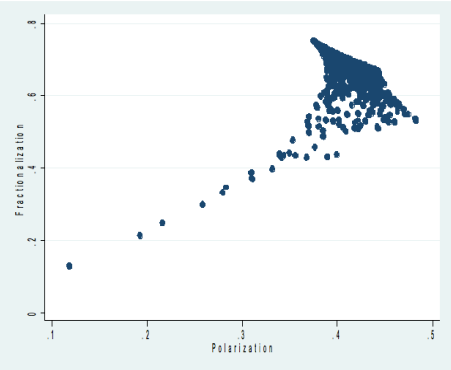
[c] As a further robustness check we compute linguistic distance as Desmet et al. (2009a) using $\delta = 0.05$ and obtain qualitatively similar results, which are available from the authors upon request.

Appendix Tables and Figures

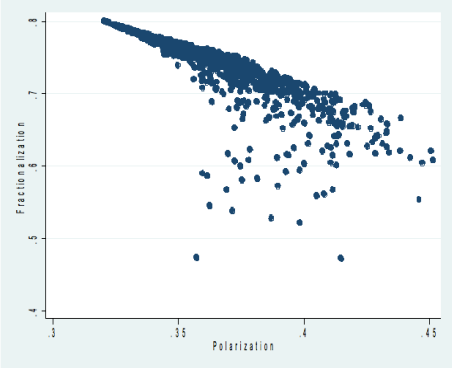
Appendix Figure 1: Fractionalization (y-axis) vs. Polarization (x-axis)



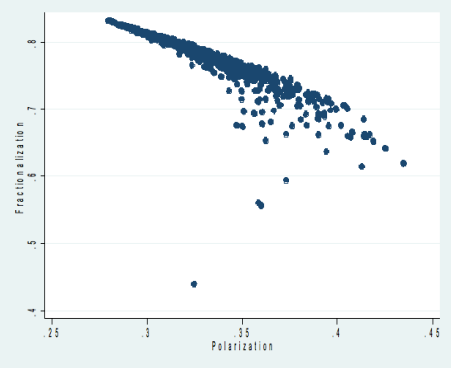
(a) Number of groups = 3



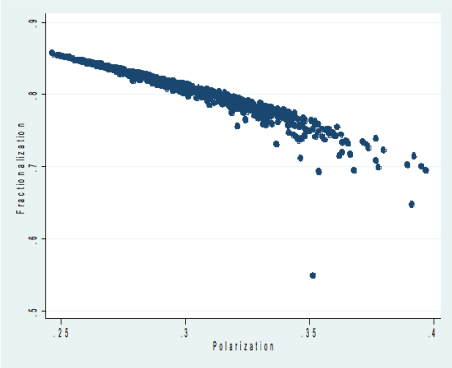
(b) Number of groups = 4



(c) Number of groups = 5



(d) Number of groups = 6



(e) Number of groups = 7

Appendix Table I
The Impact of Polarization and Fractionalization on Output Growth: Excluding South

	Output p.c. (log-change) 1870 - 1920			
	(1) LS	(2) IV (A)	(3) IV (B)	(4) IV (C)
Change POL 1870-1920	-1.398***	-2.045**	-2.210***	-1.450***
	(0.359)	(0.810)	(0.582)	(0.481)
Change FRAC 1870-1920	2.313***	3.337***	2.057***	2.039***
	(0.492)	(1.046)	(0.673)	(0.608)
Observations	1044	1044	1044	1044
R^2	0.617	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	4.40	31.03	6.09
Hansen-J statistic (p-val.)	-	0.90	n.a.	0.90
Anderson-Rubin Wald-Test (p-val.)	-	0.07	0.00	0.05
Endogeneity Test Statistic (p-val.)	-	0.83	0.07	0.16
State FE	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In column (1) the method of estimation is least squares. In columns (2)-(4) the method of estimation is two-stage least squares. In column (2) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In column (3) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In column (4), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)). The initial control variables (1870) for columns (1)-(4) are output per capita, population size, land concentration, the manufacturing share, rail access, the share of foreign-born and the share of Afro-Americans (estimates not reported in the table).

Appendix Table II
Nonlinear Relationship between Fractionalization and Polarization

	Change Output p.c. (in logs) 1870 - 1920			
	(1) LS	(2) IV (A)	(3) IV (B)	(4) IV (C)
Change POL 1870-1920	-0.715**	-1.643**	-2.145***	-1.610***
	(0.326)	(0.691)	(0.737)	(0.583)
Change FRAC 1870-1920	0.142	1.491	-0.702	-0.848
	(0.807)	(2.372)	(1.632)	(1.432)
Change $FRAC^2$ *1870-1920*	1.313**	0.406	2.598**	2.514**
	(0.607)	(2.292)	(1.159)	(1.035)
Observations	2068	2068	2068	2068
R^2	0.545	-	-	-
Kleibergen-Paap F-Statistic	-	3.11	63.16	11.67
Hansen-J statistic (p-val.)	-	0.81	n.a.	0.59
Anderson-Rubin Wald-Test (p-val.)	-	0.00	0.00	0.00
Endogeneity Test Statistic (p-val.)	-	0.26	0.00	0.00
State FE	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In column (1) the method of estimation is least squares. In columns (2)-(4) the method of estimation is two-stage least squares. In column (2) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In column (3) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In column (4), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)). The initial control variables (1870) for columns (1)-(4) are output per capita, population size, land concentration, the manufacturing share, rail access, the share of foreign-born and the share of Afro-Americans (estimates not reported in the table).

Appendix Table III
The Impact of Polarization and Fractionalization on Output Growth: Below vs. Above Median Manufacturing Output

	Output p.c. (log-change) 1870-1920: Columns (1)-(4) Above; Columns (5)-(8) Below the Median							
	(1) LS	(2) IV (A)	(3) IV (B)	(4) IV (C)	(5) LS	(6) IV (A)	(7) IV (B)	(8) IV (C)
Change POL 1870-1920	-1.376***	-2.273**	-2.499***	-1.398***	0.206	1.076	-2.188*	-0.749
	(0.347)	(0.946)	(0.738)	(0.522)	(0.542)	(0.707)	(1.133)	(0.657)
Change FRAC 1870-1920	2.340***	3.664***	2.760***	2.304***	-1.146	-2.069*	0.670	0.0963
	(0.508)	(1.285)	(0.809)	(0.726)	(1.002)	(1.068)	(1.154)	(0.941)
Observations	998	998	998	998	966	966	966	966
R^2	0.464	-	-	-	0.526	-	-	-
First-stage (Kleibergen-Paap) F-Statistic	-	8.71	25.31	10.31	-	21.00	9.45	16.22
Hansen-J statistic (p-val.)	-	0.57	n.a.	0.59	-	0.47	n.a.	0.06
Anderson-Rubin Wald-Test (p-val.)	-	0.00	0.00	0.00	-	0.00	0.03	0.00
Endogeneity Test Statistic (p-val.)	-	0.61	0.14	0.19	-	0.95	0.05	
State FE	yes	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. In columns (1) and (5) the method of estimation is least squares. In columns (2)-(4) and (6)-(8) the method of estimation is two-stage least squares. In columns (2) and (6) the instrumental variable is the supply-push component of immigrant inflows (IV (A)); also see Section 4, page 8 for further details. In columns (3) and (7) the instrumental variables are the initial 1870 polarization and fractionalization index (IV (B)). In columns (4) and (8), the instrumental variables are the supply-push component of immigrant inflows and the initial 1870 polarization and fractionalization index (IV (C)). The initial control variables (1870) for columns (1)-(8) are output per capita, population size, land concentration, the manufacturing share, rail access, the share of foreign-born and the share of Afro-Americans (estimates not reported in the table).