Summaries of Doctoral Dissertations

The Dissertations of James Feigenbaum, Timothy Larsen, and Elisabeth Perlman: 2016 Allan Nevins Prize Competition of the Economic History Association

ROBOTS, REBELS, AND RAILROADS

I was honored that Lee Alston asked me to be the convener for the Nevins Prize Competition this year. There were close to a dozen submissions spanning a wide range of topics in economic history and displaying a wide range of methodologies, from the quantitative to the narrative. Reading all of these dissertations was, for me, interesting, enlightening, and even inspiring. If the research in this most recent cohort of PhD dissertations is any indication, economic history has a bright future.

That being said, I was given the difficult task of selecting the finalists for the prize. In my judgment, three dissertations stood out as exceptional innovations. The three finalists were James Feigenbaum, PhD from Harvard University; Timothy Larsen, PhD from University of Colorado; and Elisabeth Perlman, PhD from Boston University. These dissertations achieve their novelty in a standard way: first successfully rooting around for historical data (or *capta*, as Deirdre McCloskey, 2013, would admonish) and second applying cutting-edge techniques to the analysis.

In my view, some of best research in economic history serves two separate, but interdependent functions. On the one hand, it can be read as a serious work of economic history, that is, history for its own sake. On the other hand, it can be read as confronting deep questions of general economics in which the history is taken seriously. These three dissertations accomplish both functions. I provide some brief motivation and introduction for these excellent dissertations. More detailed summaries by the authors themselves follow this article.

ROBOTS AND RECORD LINKAGE: JAMES FEIGENBAUM

James Feigenbaum's dissertation is entitled "Essays on Intergenerational Mobility and Inequality in Economic History." His work bolsters a renewed attention to the historical measurement of intergenerational mobility. This renewal is in part opportunistic; recent years have seen the increasing availability of machine-readable, historical microdata, much of which was collected for genealogy rather than social science. One sweeping example is Gregory Clark's (2014) clever use of surnames as a way of measuring (very high) persistence across generations by patrilineal descent. While the surname-based methodology answers a question that is at least somewhat distinct from the traditional father/son model of intergenerational mobility, the method has the advantage that samples can be constructed without having to track families directly. But to answer the traditional question, we need to link parents to their adult children.

For example, Jason Long and Joseph Ferrie (2013) construct linked samples across U.S. censuses to compare occupations between fathers and their linked sons. In my own work with Ferrie (Bleakley and Ferrie 2016), we link census observations to track families exposed to a large, random wealth shock in one of the antebellum Georgia land lotteries. Longitudinal linkage in these and in various other studies is based on the pioneering work of Ferrie (1996). Such linkage is done on a somewhat *ad hoc* basis, however, with a limited set of variables and rules used to match observations across datasets.

As we have heard in popular culture, the robots are coming to take our jobs. The machine-learning approach of Feigenbaum's dissertation is a step in that direction, thus currying him the favor of our new robot overlords. As a bit of background, first consider the trade-off faced by the prior literature. An ideal approach to linkage might involve an army of trained genealogists who know how to find people in historical records and understand how to recognize a spurious match. This has been the approach, approximately, of the Union Army Project (www.uadata.org) led by the late Robert Fogel and now by Dora Costa, which has a wealth of variables, but comparatively few observations. Attempts to link across publicly available Census data suffer from the opposite problem: the universe of people, but few variables. Hiring relatively expensive genealogists would be prohibitively costly when forming a large-scale sample. That said, might there be a hybrid strategy that is cheaply scalable yet incorporates aspects of human judgment in the myriad of ways that cannot be codified easily into a small set of rules?

One of the chapters of his dissertation takes on the linkage problem using tools of machine learning. The method starts with a small subsample of observations linked as carefully as possible by knowledgeable humans. The resulting links are used to train an algorithm about which combinations of characteristics are the most likely to be the correct match. The algorithm can then construct a bilateral score (a proxy for a match probability) between observations in the two data sets. Yet there are multitudinous ways in which observations could differ, and the machine-learning algorithms have various tricks for (a) limiting the set of variables used to construct scores and (b) avoiding an over-fitting problem using cross-sample validation. Feigenbaum describes various methods in a way that is clear and accessible to applied economists. He then implements the strategies by linking several early-twentieth-century datasets. The results suggest that his method can be a useful tool for constructing large longitudinal samples on a shoestring. (But also see subsequent work by Martha Bailey, Morgan Henderson, and Catherine Massey (2016), who show that humans can, and probably should, be more in the loop for linkage.)

The balance of his dissertation takes the linked samples and applies them to questions of intergenerational mobility. Chapter 2 addresses itself to the bread-and-butter question of estimating intergenerational elasticities. Chapter 1 asks whether intergenerational correlations were affected by the intensity of the Great Depression. Because his linkage procedure can efficiently generate a large, linked sample, he can precisely estimate differences in intergenerational coefficients across subsamples, in this case based on whether the father lived in a city that saw greater declines in employment in that period. These chapters are strong contributions to this literature by providing (a) the first large-sample estimates on income mobility from that period and (b) more evidence on the contextual determinants of intergenerational correlations.

REBELS, RENEWAL, AND RECRUDESCENCE: TIMOTHY LARSEN

Timothy Larsen's dissertation bears the title "Confederate Deaths and the Development of the American South" and investigates the persistence of the War's demographic impact at the county level. The death toll in the U.S. Civil War was estimated to be far higher than that from other conflicts in this nation's history. Moreover, earlier estimates substantially understate both the immediate body count and the short-term demographic impact of the War, according to recent work by David Hacker (2011), who used comparisons of cohort sizes in various censuses to estimate regional mortality rates. Yet Hacker argues that the demographic impact was relatively short lived. Further, Paul Paskoff (2008) finds that county-level measures of property destruction during the U.S. Civil War predict little difference in the postwar capital stock. This indicates that recovery from direct damage was quick, even while the adaptation to institutional change (e.g., emancipation) was slow. The speedy recovery from material damage is consistent with other studies finding a quick bounce-back after wartime destruction, for example, Donald Davis and Jeremy Weinstein's (2002) article analyzing twentieth century Japan before and after WWII. However, it is inconsistent with arguments for path dependence in urban location, for example my discussion (Bleakley and Jeffrey Lin 2012) of population persistence at sites where rivers cross the fall line, where there should be little contemporary relevance for the obstacle to navigation or other attendant historical advantages. Clearly, we need more evidence to resolve this contradiction. Enter Mr. (now Dr.) Larsen.

With elbow grease, shoe leather, and various additional metaphors for work effort, Larsen assembles county-level estimates of military deaths in the Civil War. The collection of these records and construction of the data are presented in his second chapter. Genealogists had already digitized some records, and Larsen coded the rest from microfilm curated by the Family History Library. All told, he is able to construct estimates for each county in 8 of the 11 states that attempted secession. These estimates are interesting in and of themselves, but also a valuable input to assessing the impact of demographic shocks. Civil War historian Shelby Foote noted that, because regiments were raised at a very local level, "one of those regiments can get in a very tight spot in a particular battle, like in the cornfield at Sharpsburg, and the news may be that there are no more young men in that town [where the regiment was organized] … they're all dead," (quoted on p. 3 of Larsen's dissertation). This heterogeneity across counties in war death rates allows Larsen to compare otherwise similar areas that get struck by very different shocks. His data include many neighboring counties with markedly different war mortality, for example. Further he shows that his measure of the death toll is conditionally uncorrelated with various antebellum characteristics of the counties.

The remainder of the dissertation makes fruitful use of these data (okay, *capta*, again) by comparing the evolution of outcomes in counties with high versus low death rates. In Chapter 3, he examines population and finds the high- and low-war-death counties tend to converge in the decade or so after the War. But the story does not end there; instead, he shows that high-mortality counties maintained their postwar growth boost for a century after the War. He argues that the shock disrupted those counties more and thus set them on an entirely different growth trajectory. He notes that part of this disruption was to shift the racial balance of these counties: blacks disproportionately filled the demographic void in high-mortality counties. In Chapter 4, Larsen considers

how this latter change affected race relations. He finds that the greater demographic shock (perhaps with the initial labor scarcity) depressed racial violence (e.g., lynchings) and facilitated black political participation in the postwar nineteenth century. While this reflects an unexpectedly felicitous response to devastation and labor scarcity, its character changed (thus, a "recrudescence" of racial conflict). In the twentieth century, high-damage counties saw relatively lower black political participation and relatively higher racially motivated violence. As Larsen argues, this cuts against the idea that group animosity is typically a highly persistent feature of an area or culture. Instead, it apparently can be influenced by contextual factors as well, to the point of generating rank reversals.

RAILROADS, R&D, AND RFD: ELISABETH PERLMAN

Elisabeth Perlman's dissertation is "Connecting the Periphery: Three Papers on the Developments caused by Spreading Transportation and Information Networks in the Nineteenth-Century United States." In the time that it took you to read this lengthy title, a tweet or video can "go viral" on today's social media. But information diffusion was a fair bit slower in the nineteenth century. Perlman's dissertation considers the impact of nineteenth-century advances in transportation and communication technology on the diffusion of information. Economic historians have studied transport improvements for some time. Examples abound, but I will give only two. Fogel (1964) famously computed the "social savings" of the railroad. More recently, Jeremy Atack et al. (2010) considered the effect of the railroad on urbanization, and, in the process, generated a boatload (wagonload?) of geocoded railroad *capta* that have enabled work in this area by numerous other scholars.

Perlman asks if the arrival of the railroad spurs innovation. Following a large literature, she uses patents to measure innovation. She uses a variety of sources to construct geocoded patent data (more *capta*, actually) for the nineteenth century. It is easy to spend a lot of time just looking at the summary statistics. These data lend themselves to maps, where you can easily see the westward expansion of the country and also that the South becomes something of a "patent desert." Another fascinating fact that emerges is the secular decline in the spatial concentration (measured by a Herfindahl-Hirschman index) over the long nineteenth century. This contrasts with an inverse-U-shaped relationship in the change of population concentration over the same time. But the headline result is that patenting per capita increases with the arrival of transportation access, be it via rail or canal.

She then turns to evaluating a standard economic-geography mechanism for these results. Specifically, she considers the idea that transportation gives access to a larger market, and that this higher scale gives an added incentive to the high-fixed-cost activity of innovation. She uses a gravity model to construct a measure of market access. When the railroad arrives, it does not affect everyplace in the same way. Instead, it depends on how the new transportation changes your effective distance (a sort of weighted transport cost) to the rest of the world. She finds that market access explains very little of the effect of transportation on patenting. It seems that the story behind her result is not about scale, at least in the usual economic-geography sense, but rather about something else that happens to a place when it becomes more robustly connected to the rest of the world.

Perlman's next two chapters continue her analysis of information diffusion. The second chapter adds an additional layer of technical sophistication by delving in to the raw text of the patents with automated techniques. Again, the descriptive statistics are themselves fascinating. She can track the diffusion of new words appearing in patents, such as the diffusion of the word "vulcanization." Nevertheless, she finds that transportation access is a less important driver of being on the cutting edge of new techniques, at least as far as she can measure them with new words. Her third chapter sets off in a different direction by looking at the effect of Rural Free Delivery (RFD) mail on voting in the early twentieth century. This should affect information diffusion, and indeed she shows that national politicians adjust their positions to be closer to their rural constituents.

HOYT BLEAKLEY, *University of Michigan*

REFERENCES

Atack, Jeremy, Fred Bateman, Michael Haines, et al. "Did Railroads Induce or Follow Economic Growth?: Urbanization and Population Growth in the American Midwest, 1850–1860." *Social Science History* 34, no. 2 (2010): 171–97.

Bailey, Martha, Morgan Henderson, and Catherine Massey. "How Well Do Automated Linking Methods Perform? Evidence from the LIFE-M Project." Unpublished Manuscript, University of Michigan, 2016.

Bleakley, Hoyt, and Joseph Ferrie. "Shocking Behavior: Random Wealth in Antebellum Georgia and Human Capital Across Generations." *Quarterly Journal of Economics* 131, no. 3 (2016): 1455–95.

Bleakley, Hoyt, and Jeffrey Lin. "Portage and Path Dependence." *Quarterly Journal of Economics* 127, no. 2 (2012): 587–644.

Clark, Gregory. *The Son Also Rises: Surnames and the History of Social Mobility.* Princeton: Princeton University Press, 2014.

Davis, Donald, and David Weinstein. "Bones, Bombs, and Break Points: The Geography of Economic Activity." *American Economic Review* 92, no. 5 (2002): 1269–89.

Ferrie, Joseph. "A New Sample of Males Linked from the Public-Use-Microdata-Sample of the 1850 US Federal Census of Population to the 1860 US Federal Census Manuscript Schedules." *Historical Methods* 29 (1996): 141–56.

Fogel, Robert W. *Railroads and American Economic Growth: Essays in Econometric History.* Baltimore: Johns Hopkins University Press, 1964.

Hacker, J. David. "A Census-Based Count of the Civil War Dead." *Civil War History* 57, no. 4 (2011): 306–47.

Long, Jason, and Joseph Ferrie. "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *American Economic Review* 103, no. 4 (2013): 1109–37.

McCloskey, Deirdre. "The Captivating Fogel." *Capital Ideas.* (Magazine of The University of Chicago, Booth School of Business). Winter 2013. Also available online at http://deirdremccloskey.com/editorials/fogel.php.

Paskoff, Paul F. "Measures of War: A Quantitative Examination of the Civil War's Destructiveness in the Confederacy." *Civil War History* 54, no. 1 (2008): 35–62.

Essays on Intergenerational Mobility and Inequality in Economic History

This dissertation explores the historical roots of economic inequality and mobility and looks to the past to expand our understanding of contemporary economic questions. In particular, I focus on intergenerational mobility in the early twentieth century United States. In the first chapter, I examine the effects of economic conditions on mobility. In the second chapter, I compare mobility rates across time. In the third chapter, I document the new tools used in the previous chapters to construct linked longitudinal micro-data across generations.

The first chapter asks whether economic downturns expand or contract intergenerational economic mobility. The effects of macroeconomic conditions on the stakes of the lottery of birth are unclear. A significant disruption of the economy may diminish or even render irrelevant inequities of opportunity bestowed by the previous generation, decoupling the fates of children from their parents. Alternatively, poorer families may be less able to endure a downturn, and children who might have climbed the income ladder in normal times—perhaps with more education or by making savvier migration choices—would instead emerge from a crisis no better off than their parents. To answer this question, I estimate the effects on intergenerational mobility of the largest economic cataclysm in American history, the Great Depression. The Depression presents an ideal natural experiment for studying the impact of a downturn on the transmission of economic status for three primary reasons. First, with a quarter of the labor force unemployed, the magnitude of the Depression dwarfs the recessions of the postwar era; if downturns do alter mobility rates, this should be most observable in the largest one. Second, fortunes large and small were destroyed in the 1929 stock market crash as well as in many local real estate crises and bank failures across the United States, creating geographic variation that enables me to compare intergenerational mobility across cities affected unevenly by the Depression. Third, the passage of time allows me to observe the downturn-exposed children as adults in 1940, which is not yet possible for the children of the recent Great Recession.

I measure mobility in two historical datasets that I construct, linking parents before the Depression to their children as young adults in 1940. Historical microdata reporting earnings as well as names and ages—necessary to link records between sources—are exceedingly rare before 1940.[1] I add to the stock of these datasets by digitizing and transcribing the 1918–1919 Bureau of Labor Statistics (BLS) cost of living survey that includes names, ages, addresses, income, and occupation for 12,817 families in 99 cities. I link the parents in the survey ahead to their children in 1940, creating a new dataset of intergenerational earnings. I supplement this data by building a second

James Feigenbaum, Assistant Professor, Department of Economics, Boston University, Boston, MA. E-mail: james.feigenbaum@gmail.com. This dissertation was completed under the supervision of Claudia Goldin (chair), Lawrence Katz, Richard Hornbeck, and Edward Glaeser at Harvard University.

[1] In the United States, there are two such sources: the 1915 Iowa State Census, which includes records for only the state of Iowa (Goldin and Katz 2000), and published lists of personal income tax payers in New York City in 1923 and 1924 (Marcin 2014). The 1940 Federal Census is the first census to include data on earnings and completed years of education. The original manuscripts with names and addresses were released in 2012 following the standard 72-year privacy period.

linked sample of parents and children containing individual occupations and average occupational earnings, drawn from the 1920 and 1940 U.S. Censuses of Population. I document the machine learning approach to record linkage in the third chapter of the dissertation, discussed later. I combine the linked data with measures of local variation in Great Depression severity based on retail sales from the work of Price V. Fishback, Shawn Kantor, and John Joseph Wallis (2003).

I find that economic mobility was lower in cities with more severe downturns during the Depression: both the 1940 earnings and occupations of sons growing up in these cities are more closely linked to their father's outcomes than are those of sons in less negatively affected cities. I measure relative intergenerational mobility in three ways: the elasticity of the son's earnings with respect to the father's earnings; the coefficient from a regression of the son's position in the earnings distribution in 1940 on the father's position in 1920; and the elasticity of the son's occupation score to the father's occupation score.[2] All three measures show similar reductions in mobility caused by the Great Depression. The effects are of similar magnitude for all sons in my sample, regardless of whether they were in grade school in 1929, of high school age, or already in the labor force when the Depression hit.[3]

To explore whether the Depression effects on mobility are causal—and not driven by pre-existing city-level differences in mobility that happen to correlate with Depression severity—I perform a parallel analysis on intergenerational mobility from 1900 to 1920. Implicitly, this is a differences-in-differences framework, comparing sons in cities before and after the Depression that experienced Depression downturns of varying magnitudes. For this earlier generation, I find no ex ante differences in mobility between cities that later experienced larger or smaller Depression downturns. This result suggests that it is unlikely that low levels of mobility drove local Depression severity, or that mobility and Depression severity are correlated outside the Depression generation. I also show that Depression severity does not predict intergenerational mobility in the late twentieth century.

How did the Depression decrease mobility in cities that experienced more severe shocks? Migration is a key mechanism. Local Depression severity drove out-migration, as sons fled distressed cities for better opportunities elsewhere. However, not all sons migrated at the same rates or to the same places. Rather, migration varied by father's earnings: the sons of richer fathers were able to make better migration decisions, moving to cities and regions that had suffered less severe Depression downturns, reducing estimated economic mobility. Formal education plays no role in explaining my estimated effects; while sons from cities with more severe downturns did accumulate more years of schooling than sons from other cities, the Depression did not affect the link between a father's earnings and a son's education. I also find that the Depression increased income inequality in cities, a change that may have made climbing the ladder of economic status more difficult. Finally, I show that local variation in New Deal spending did not affect mobility.

The estimated effects of the Depression on mobility are economically large. The differences in intergenerational mobility rates between sons in cities with more severe

[2] I focus on relative intergenerational mobility, estimated via the slope of a regression of the sons' outcomes on their fathers' outcomes, rather than absolute mobility, which indicates the expected position of a child born to a parent at a given place in the earnings distribution.

[3] I estimate the relative effect of the Great Depression, exploiting variation in severity across cities. I refer to this local, differential cross-city impact of the Depression as the Great Depression effect. However, I am unable to directly assess the aggregate effect of the Depression on intergenerational mobility.

and less severe downturns are comparable to the differences in mobility between the United States and Sweden today, the least and most economically mobile countries in the Organisation for Economic Co-operation and Development (OECD) (Corak 2013). The differences I find are also comparable to the differences in mobility between Charlotte, NC and Salt Lake City, UT, two American cities that are currently at opposite ends of the mobility spectrum (Chetty et al. 2014a). The Depression calcified the mobility ladder for the generation of children unlucky enough to be born in cities with especially severe downturns.

In the second chapter, I ask whether intergenerational economic mobility was higher in the early twentieth century in the United States than today. Comparing mobility over time is complicated by different measures of mobility across studies; often the chosen measures are influenced and constrained by the available data. In historical work on intergenerational mobility, income or earnings data is rare, and so occupations and (more recently) names are the proxies for status used.[4] In contemporary data, scholars often calculate mobility with earnings and education.[5] But recent studies on trends in intergenerational mobility are unable to trace income mobility before the 1980s (Chetty et al. 2014b; Lee and Solon 2009).

I take advantage of two historical data sources that enable me to measure mobility in many different ways—including intergenerational elasticity of earnings, rank-rank persistence, occupational score elasticity, occupational transitions, education persistence, and imputed status based on both given names and rare surnames. Because I can estimate these various mobility measures within the same source, I shed light on how well these various measures agree with one another, at least for the early twentieth century. I match fathers from the Claudia Goldin and Lawrence F. Katz (2000) sample of the Iowa State Census of 1915 to their sons in the 1940 Federal Census, the first state and federal censuses with data on income and years of education.[6]

The different measures of intergenerational mobility are quite consistent. I find generally high rates of mobility across all measures. These measures are also internally consistent: I estimate more mobility for the sons of urban Iowa than for the sons of rural Iowa, as well as more mobility for the grandsons of the foreign born than native born grandsons.[7]

---

[4] While the United States federal census began collecting information on respondents' occupations in 1850, the census did not include data on either years of educational attainment or annual income until 1940.

[5] Even with earnings data, sharp disagreements remain over whether to estimate intergenerational elasticities or rank-rank mobility and the effects of the observation of multiple years of earnings, possibly at the household or individual level, enabled by administrative data.

[6] My study builds on the earlier mobility work of Parman (2011), who also draws on the 1915 Iowa State Census to measure intergenerational mobility, linking men in the 1915 Iowa census back to their households in the 1900 census and then finding their fathers and sons in Iowa in 1915.

[7] I use grandparent nativity rather than son or father nativity because the subsamples are more even: nearly one-third of my sample have four foreign-born grandparents, another one-third have four native-born grandparents, and the rest 1, 2, or 3. A much smaller share have foreign born parents and only a handful of the sons are foreign born themselves. I am unable to estimate mobility across race in my sample because Iowa's population in 1915 was nearly all white. Hertz (2009) documents strong disparities in mobility between blacks and whites, arguing that American immobility in the recent period is driven by extremely high persistence of outcomes among African Americans.

Was there more mobility historically than today? I find a lower intergenerational income elasticity during the first half of the twentieth century in the United States than studies find in the second half of the century. I also measure intergenerational mobility using the rank-rank parameter and similarly find more mobility historically. The results for education, occupational mobility—measured either using transition matrices or occupation score—and name-based mobility all point in the same direction. But such differences between contemporary and historical mobility could be spurious. For one, my sample is not a random draw of the American population. By population density, urbanization, and the share foreign born, Iowa is nearly the median state. On other dimensions, Iowa in 1915 is an outlier: it was also almost universally white, invested more and earlier in education than other states, and had relatively low levels of inequality. In addition, measurement error, either due to the historical nature of my data, the difficulty of creating longitudinal linked samples, or the single year I am able to observe fathers and sons in the census, may push me to find excess mobility. I address these concerns in a few ways. In the earnings data, I show that the estimated differences between contemporary and historical mobility remain after adjusting the contemporary sample to mirror the historical sample in measurement noise and demographic and geographic composition. The agreement across many measures of mobility, some based on outcomes like education and occupation that should be less noisy than a single year of earnings, also sharpens the comparative result. Ultimately, my results do not prove there is less mobility today than in the past, but, taken along with other evidence comparing mobility over the twentieth century (Parman 2011; Long and Ferrie 2013; Olivetti and Paserman 2015), they strengthen the argument that there was more mobility in the early twentieth century.

The final chapter details the machine learning approach to record linkage used to construct the samples I analyze in the first two chapters. Newly digitized complete historical census records have brought big data to economic history. Linking individuals over time and between databases has opened up avenues for research into intergenerational mobility, the long-run effects of early life conditions, assimilation, discrimination, and the returns to education. To take advantage of these new research opportunities, scholars need to be able to accurately and efficiently match historical records and produce an unbiased dataset of links for analysis. The procedure I propose applies insights from machine learning classification and text comparison to record linkage of historical data.[8]

Linking historical data—without unique identification numbers—is difficult and imprecise, relying on demographic information like name, age, and place of birth. However, these variables may be mismeasured, due to transcription errors, spelling mistakes, name changes, or name shortening. Manual linking by a trained researcher yields accurate and comprehensive matches, but at the cost of time and replicability. Prior algorithmic approaches have been developed in the historical literature, but their rigid rules are often quite inefficient—many records go unmatched—and inaccurate in the face of messy historical data. My technique uses supervised learning to train an algorithm to replicate the process of manually matching individual records across

---

[8] The use of machine learning techniques in the economics literature has increased recently both in traditional prediction tasks (Kleinberg et al. 2015) and in causal policy evaluation (Athey 2015).

sources. I am thus able to increase the speed, accuracy, and consistency of creating historical linked samples.

I detail the specifics of the linking method in the chapter and give a brief overview here. I begin by cross matching two census-like datasets and extracting a wide subset of possible matches for each record. I then build a training dataset on a small share of these possible links, manually identifying whether a given record pair is a link or not. This is the data used to tune the matching algorithm. For every record pair, I generate a large set of features—string distance in first and last name, difference in year of birth, Soundex indicators, agreement on middle initial, total possible matches for a given record, etc. When a researcher is making matches manually, each of these features has some weight: differences in last name are penalized more than differences in first names or string differences early in names are penalized more than string differences at the end of names. However, these weights are only implicit, and a researcher might struggle to write them. The algorithm, observing the features and the match or not-match outcome, attempts to minimize both false positives and false negatives and estimates the weights explicitly, learning the rules used by a well-trained and consistent researcher.

<div align="right">JAMES FEIGENBAUM, *Boston University*</div>

REFERENCES

Athey, Susan. "Machine Learning and Causal Inference for Policy Evaluation." In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Conference held in Sydney, NSW, 2015.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, et al. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics* 129, no. 4 (2014a): 1553–1623.

———. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review: Papers & Proceedings* 104 (2014b): 141–47.

Corak, Miles. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives* 27, no. 3 (2013): 79–102.

Fishback, Price V., Shawn Kantor, and John Joseph Wallis. "Can the New Deal's Three Rs Be Rehabilitated? A Program-By-Program, County-By-County Analysis." *Explorations in Economic History* 40, no. 3 (2003): 278–307.

Goldin, Claudia, and Lawrence F. Katz. "Education and Income in the Early Twentieth Century: Evidence from the Prairies." *Journal of Economic History* 60, no. 3 (2000): 782–818.

Hertz, Tom. "Rags, Riches, and Race: The Intergenerational Economic Mobility of Black and White Families in the United States." In *Unequal Chances: Family Background and Economic Success*, edited by Samuel Bowles, Herbert Gintis, and Melissa Osborne Groves, 165–91. Princeton: Princeton University Press, 2009.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, et al. "Prediction Policy Problems." *American Economic Review* 105, no. 5 (2015): 491–95.

Lee, Chul-In, and Gary Solon. "Trends in Intergenerational Income Mobility." *Review of Economics and Statistics* 91, no. 4 (2009): 766–72.

Long, Jason, and Joseph P. Ferrie. "Intergenerational Occupational Mobility in Great Britain and the United States Since 1850." *American Economic Review* 103, no. 4 (2013): 1109–37.

Marcin, Daniel. "Essays on the Revenue Act of 1925." Diss., University of Michigan, 2014.

Olivetti, Claudia, and M. Daniele Paserman. "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850–1940." *American Economic Review* 105 (2015): 2695–724.

Parman, John M. "American Mobility and the Expansion of Public Education." *Journal of Economic History* 71, no. 1 (2011): 105–32.

Confederate Deaths and the Development of the U.S. South

The American Civil War casts a long shadow over Southern history. For a century afterwards, and probably much longer, history impacted Southern life in ways unknown to other parts of the nation. Toynbee, from Britain, observed:

> "History is something unpleasant that happens to other people. We are comfortably outside all that. If I had been a small boy in New York in 1897 I should have felt the same. Of course, if I had been a small boy in 1897 in the Southern part of the United States, I should not have felt the same; I should then have known from my parents that history had happened to my people in my part of the world." (Woodward 1968, p. 189–90)

In this dissertation, I study changes in the demography, labor markets, and racial systems of Southern counties caused by deaths in the Confederate Army. Population loss was enormous, with the South losing more men in percentage terms than any nation lost in WWI (McPherson 1996). Outside of ending the slave system, the war's death toll was likely its largest impact on later generations. Still, Southern deaths in the Civil War have received little formal treatment, especially at the local level (see Hacker (2001) for an analysis of Civil War deaths on aggregate demographic measures in the United States, Faust (2008) for a social history of Civil War deaths). Consequently, little is known of the impact of these deaths in the South, a region which endured a regrettable century from surrender (1865) to the Civil Rights Era.

To estimate the effects of military deaths, I construct county-level estimates of deaths in the Confederate army. These are the first local-level estimates of losses in the Civil War, North or South. Importantly, the death rates were determined by where Confederate units, which were geographically homogenous, were sent to fight for strategic reasons. The death rates are uncorrelated with counties' pre-war characteristics, allowing for causal inference. I then use the variation in these death rates to estimate

Tim Larsen, Assistant Professor, Department of Economics, Berry College, Mount Berry, GA. E-mail: tlarsen@berry.edu. Dissertation completed at the University of Colorado Boulder under the direction of Murat Iyigun, Lee Alston, and Ann M. Carlos.

the war's impact on the Southern population and economy, with particular emphasis on outcomes for African Americans, who did not die in the war but were greatly affected by the demographic shock the war created.

In response to population loss, there was heavy migration into counties with higher death rates in the Civil War. County populations recovered their relative standing within 15 years after the war. Black migration was especially large, and black populations remain relatively larger in those counties today, 150 years later. Military deaths particularly affected labor markets, as nearly all the deaths were white men in their prime working years. Labor scarcity in some counties led to better racial treatment for African Americans through 1900, during a time when the Southern racial system was not yet consolidated. In the very long run, however, relative discrimination was much worse in these counties, which had large black populations but whose labor scarcity had long since vanished. This gives evidence of considerable flexibility in Southern racial discrimination in response to changing incentives.

COUNTY-LEVEL DEATH RATES FOR THE CONFEDERATE ARMY

Population loss from the American Civil War was enormous, with more than 750,000 deaths (Hacker, 2011), more than all other American wars combined. The South bore a disproportionate number of the fatalities, losing at least one-in-four white men of military age (Vinovskis 1989). Overall the former Confederate states lost 4 percent of their population—by comparison, J. David Singer and Melvin Small (1972) found that in 93 wars between 1816 and 1965, there were few cases where military deaths exceeded even 2 percent of the prewar population. Despite the magnitude of the losses, relatively little is known about how many men died at the state and local levels.

From a variety of sources, I have assembled county-level estimates of deaths in the American Civil War for 8 of the 11 former Confederate States (Virginia, South Carolina, Georgia, Florida, Alabama, Mississippi, Louisiana, and Arkansas), accounting for 70 percent of the Confederacy's population. Using military records, I count deaths by Confederate company (units of roughly 100 men). Confederate companies were raised at the local level, generally composed of men from the same city or county—with this I am then able to map companies' deaths back to their home counties. I construct death rates as the percent of the military-aged population (white males aged 15–39 in the 1860 Census) from each county who died in the war.

The distribution of death rates across counties has a particularly wide variance due to the recruitment pattern. Men from neighboring counties were often sent to different theaters, and even assignments within battles led to large disparities in death rates. For example, Company F of the 26th Regiment North Carolina Infantry experienced 100 percent casualties (killed, captured, or wounded) at Gettysburg (McGee 2014). Nearly all of those men are documented as coming from Caldwell County, which was sparsely populated, before the war.

Generals made troop assignments for strategic reasons, so the death rates are unrelated to counties' pre-war characteristics. Suggestive evidence of this can be seen as the death rates are uncorrelated with observable characteristics from the 1860 Censuses of Population, Agriculture, and Manufacturing. Confederate armies were also very large, often comprised of units from many states and hundreds of counties—men from a single county would be unlikely to have affected generals' decision-making for armies this large. Furthermore, many battles in the Civil War were unplanned, simply occurring as

the armies found each other. Gettysburg, for example, began as Confederate soldiers descended on the city in search of a rumored cache of shoes, instead encountering a Union cavalry unit. Since counties' death rates were randomly determined, I interpret all relationships between deaths and outcome variables causally.

Using data on the battles in which each company fought, I also construct predicted death rates for each county. The predicted death rates should vary only due to the generals' troop assignments and not due to any unobserved factors uncorrelated with those decisions, such as a county's pre-war health conditions. All of the estimated effects of counties' death rates are robust to using the predicted death rates in two-stage least squares regressions.

SWITCHING PATHS: CONFEDERATE DEATHS AND THE ECONOMIC GEOGRAPHY OF THE POSTBELLUM SOUTH

I first explore the effects of the war on the economic geography of the South—where Southerners chose to live for decades afterwards. Counties with more Civil War deaths caught up to neighboring areas' populations within 15 years. They then continued to grow, pushing far past convergence with neighboring counties by 1900: counties with 10 percentage-point higher death rates (roughly one standard deviation) had 9 percent larger populations in 1900, and by 1960, nearly 100 years after the war, the margin grew further to 16 percent. I find even stronger results for counties' African American populations: counties with 10 percentage-point higher death rates had 14 percent larger black populations by 1900 and 27 percent larger by 1960. Black populations in these counties remain higher than in otherwise similar counties today.

Population growth appears to have come through migration, especially black migration, to counties with higher death rates in the Civil War. The economic geography of the South was also shifted following the war as migrants increasingly went to labor-scarce counties participating in newer industries, such as resource-intensive manufacturing. They decreasingly went to labor-scarce counties that were advantaged in the pre-war economy, such as plantation counties and counties with early rail access.

LABOR SCARCITY AND DISCRIMINATION, 1865–1900

Of all the South's rebuilding efforts after the war, most attention by far has been paid to its attempts to reshape its racial system in the wake of emancipation. There is considerable argument as to how extensive discrimination was in the Postbellum period and to what extent the later rules of the Jim Crow system were already in place soon after the war. For example, in *The Strange Career of Jim Crow* (1955), C. Vann Woodward argues that there was considerable variation in discrimination and segregation in the South until 1900. He further rejects the then-prevailing notion that the Southern racial system on the eve of the Civil Rights Movement was "as it always had been."

I test whether acts of "non-market" discrimination (discrimination that occurs outside of the labor market) were affected by the labor scarcity counties experienced following the Civil War. By economic theory, competitive markets should eliminate taste-based discrimination in labor markets—I test whether they can affect discrimination more broadly in society as well. Labor scarcity led counties to have less racial violence— counties with 10 percentage-point higher death rates in the Civil War had 24 to 33 percent fewer lynchings from 1866–1900. These counties also had greater political

participation from African Americans as counties with 10 percentage-point higher death rates had 3.6–5.6 percent higher voter turnout from 1868–1900. These effects lasted for at least two decades after relative labor scarcity across counties had vanished, suggesting some shift in the counties' racial norms for that time period.

DISCRIMINATION, 1900 TO THE CIVIL RIGHTS ERA

In the very long run (100 years after the Civil War), however, racial outcomes saw an unfortunate reversal. Voting from 1920–1960 and black voter registration in 1960 were much lower in previously labor-scarce counties. Racially motivated murders in from 1950–1970, including murders of civil rights workers of both races, were 55 percent more likely in counties with 10 percentage-point higher death rates in the Civil War. These counties were also more likely to refuse to integrate their schools, precipitating court actions. While the mechanism is unclear, incentives had changed in these counties, no longer labor-scarce and disproportionately African American from the migration in the Postbellum period. Similar positive correlations between the size of minority populations and discrimination are consistent across studies of twentieth century United States (Oliver and Mendelberg 2000). These reversals strongly reject the idea of that levels of discrimination were either constant or culturally determined in the U.S. South—in just 2–3 generations, the same areas that had prevented lynchings were now killing Civil Rights workers. This contrasts strongly with the findings of Nico Voigtlander and Hans-Joachim Voth (2012) and Saumitra Jha (2009) who find an extreme persistence of cultural traits, lasting several centuries. The results further support Woodward's thesis: discrimination varied widely and had changed considerably by the Civil Rights Era.

<div align="right">TIM LARSEN, Berry College</div>

REFERENCES

Faust, Drew Gilpin. *This Republic of Suffering: Death and the American Civil War*. Vintage Civil War Library. New York: Vintage, 2008.

Hacker, J. David. "The Human Cost of War: White Population in the United States, 1850–1880." *Journal of Economic History* 61, no. 2 (2001): 486–89.

———. "A Census-Based Count of the Civil War Dead." *Civil War History* 57, no. 4 (2011): 307–48.

Jha, Saumitra. "Trade, Institutions, and Ethnic Tolerance: Evidence from South Asia." *American Political Science Review* 107, no. 4 (2013): 806–32.

McGee, David H. *In Spite of the World: 26th North Carolina Regimental History*. Society for the Preservation of the 26th Regiment North Carolina Troops, Inc., 2014.

McPherson, James. *Drawn with the Sword: Reflections on the Civil War*. Oxford: Oxford University Press, 1996.

Oliver, J. Eric, and Tali Mendelberg. "Reconsidering the Environmental Determinants of White Racial Attitudes." *American Journal of Political Science* 44, no. 3 (2000): 574–89.

Singer, J. David, and Melvin Small. *The Wages of War, 1816–1965*. Beverly Hills, California: SAGE Publications, 1972.

Vinovskis, Maris A. "Have Social Historians Lost the Civil War? Some Preliminary Demographic Speculations." *Journal of American History* 76, no. 1 (1989): 34–58.

Voigtländer, Nico, and Hans-Joachim Voth. "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany." *Quarterly Journal of Economics* 127, no. 3 (2012): 1339–92.

Woodward, C. Vann. "The Irony of Southern History." *The Burden of Southern History,* Revised Edition. Baton Rouge: Louisiana State University Press, 1968.

———. *The Strange Career of Jim Crow.* Oxford: Oxford University Press, 1955.

Connecting the Periphery: Three Papers on the Developments Caused by Spreading Transportation and Information Networks in the Nineteenth Century United States

The world was very poorly connected at the beginning of the nineteenth century, but by its end the movement of people, goods, and information among places had increased dramatically. Before railroads, waterways were by far the most efficient way to transport goods. Moving goods over land, even on the best roads, was extremely costly.[1] By the 1860s, the telegraph had become the fastest communication method, and physical packages moved overland by train or over water via steamship. These changes made transportation and communication faster, cheaper, and safer, effectively reducing the distance between locations. This reduction in distance was even more dramatic on the periphery of the transportation and communication network. This promoted economic growth by linking far-flung factor and product markets, encouraging the exploitation of regional comparative advantage (Fogel 1964; Atack, Haines, and Margo 2011). Railroads also changed the character of the areas around them. They provided loci for new towns—increasing urbanization (Atack et al. 2010), attracting banks (Atack, Jaremski, and Rousseau 2014), and encouraging speculators to plat towns (Hudson 1985).

This dissertation focuses on how transportation and information networks change the geographic distribution of economic activity and explores their possible democratizing effect. The first and second chapters examine the geographic distribution of patenting in the nineteenth century United States. The third explores the early twentieth century impact of Rural Free Delivery (RFD) on voting behavior. Taken together, this work reinforces our understanding of how transportation and information networks connect rural areas to emerging national markets.

Elisabeth Ruth Perlman, Economist, Center for Economic Studies, U.S. Census Bureau. E-mail: elisabeth.perlman@alumni.carleton.edu. This dissertation was completed under the supervision of Robert A. Margo (chair), M. Daniele Paserman, and Carola Frydman at Boston University.

[1] It has become commonplace to note that it cost about the same amount to ship goods between London and Boston as to travel 30 miles over land in the United States, or about the distance from downtown Boston to Concord, MA (Howe 2007).

Since Alfred Marshall (1890), economists have believed that location matters for innovative activity, as co-location facilitates the transfer of innovative ideas (Jaffe, Trajtenberg, and Henderson 1993). In Chapter 1 of this dissertation, I investigate how the nineteenth century "transportation revolution" relocated innovative activity in the United States. I first document that the expansion of improved transportation within a county increased innovation, as measured by patenting activity. I then test the hypothesis, formulated by Kenneth Sokoloff (1988), that improving transportation networks incentivized innovation by facilitating access to larger markets. I find modest evidence that market access explains the increase in patent activity, but other variables correlated with transportation access seem to explain most of the relationship.

Using a newly collected panel dataset at the county level spanning the nineteenth century United States, I find a robust, statistically significant, and positive effect of local access on patenting. This estimate suggests that the spread of transportation caused 8 percent of the increase in patenting over the nineteenth century. Because the most developed places, which already had significant transportation access, led the massive mid-century increase in patenting, increased transportation access cannot be the sole explanation. However, transportation had a significant effect for those counties not previously well connected; when I restrict my attention to only counties not saturated by transportation in 1850, this estimate doubles. In addition to a positive main effect, increased transportation is associated with decreased concentration of patenting, particularity for more rural places. To address concerns that the documented relationship between transportation access and patenting may be driven by endogeneity in the construction of new transportation, I use straight lines drawn between the prosperous places in 1830 as an instrument for transportation access. This IV specification gives much larger point estimates (they imply that more than 20 percent of the increase in patenting was due to transportation), but the standard error is such that I cannot reject the hypothesis that the IV estimates and ordinary least squares (OLS) estimates are the same.

In order to more directly test the effect of increases in the ability of an area to access larger markets, I calculate a measure of market access inspired by the methodology developed in Dave Donaldson and Richard Hornbeck (2016), which itself builds on earlier studies on this topic (Harris 1954; Gutberlet 2014). This estimate is the sum of the population of all counties in the United States; each county's contribution to this sum is weighted by the cost of moving a ton of goods from the observation county to it. The correlation of my estimate of market access with patenting per capita is not robust to the inclusion of controls. In particular, the inclusion of the lagged percent of county's improved farmland absorbs the relevant variation. Also, when both market access and local transportation access are included in the specification, local transportation access retains a positive, precisely estimated coefficient, but market access does not. It is also notable that when the percent of a county's area within some distance from transportation is used as a local access measure, the measures calculated using shorter distances relate more closely to increases in patenting. This suggests the impact of local transportation access on the number of patents comes through localized changes within a county.

Using a simple count of patents, as my first chapter does, ignores their content. Not every patent represents the same contribution to the movement of the technology

frontier, and some patents represent more important innovative contributions than others.[2] The second chapter focuses on the text of patent grants as its basic object of study, again asking how the growing transportation network changed the spatial distribution of innovation. Using the patent grants' contents provides a more nuanced study of the innovations patents represent. I propose a novel way to study technology diffusion, investigating how transportation changes information absorption.

In addition to moving goods, transportation facilitates greater communication between areas. Both railroads and canals moved passengers, allowing for easier, faster, and shorter trips between locations.[3] This greater exchange of ideas may change where innovation occurs by helping more areas learn about new technologies. This would not be unprecedented. Several studies have examined access to a different communication medium, the Internet, on the location of innovation, finding that "diffusion of the Internet worked against the trend toward increasing geographic concentration of inventive activity" (Forman, Goldfarb, and Greenstein 2014), and that an increase in communication seemed to allow greater task specialization (Agrawal and Goldfarb, 2008).

Economists have begun to use automated text analysis tools. Using modern patents, Nathan Goldschlag (2015) looks for evidence that patents help promote diffusion by searching future patents and scientific articles for text substantially similar to the language used in a patent. Goldschlag tested this technique on scientific articles known to be influential and found knowledge transmission, but found no evidence of technology transmission when using the same technique on modern patents.

To investigate the importance of information about new technologies in the nineteenth century, I use the text of patents as documentation of the movement of idea-use over time and space. I construct a measure of how many new words, from a list of words relating to new technologies, appear in the patent record of any given county. If in a particular county one of these words is observed one year after it is first used in a patent anywhere, then this county is measured as receiving new words at the rate of one new word a year. This measures the use of new technologies in patents originating from a county, telling us how quickly novel technologies become present in inventors' work. In contrast to the strong relationship between local transportation access and patents per capita, increases in local transportation access seem to have no effect on the novelty of innovation (a measure of quality).

The spread of transportation does not have a uniform effect on counties' speed of mentioning new technologies. Increased local transportation access has no net overall effect, while increasing market access increases the speed of mentioning new technologies more for more developed places. The net effect of transportation network expansion is this increase of advantage of more developed places. However, the level of patent system participation increases in every location impacted by the transportation network's spread.

---

[2] While people have addressed this in a number of ways when working with modern patents, the most popular is to use patent citations. However, patents did not start citing each other until the twentieth century.

[3] Any landing along a canal or a railroad stop was also a place where mail was exchanged. The United States had a very developed mail system early in its history; providing for the mail is one of the responsibilities laid out for the federal government in the Constitution. Mail service was adaptive, taking early advantage of new modes of transportation.

To understand this pattern in more detail, I examine a sample of patents in newly connected counties. This examination suggests that these places patented objects related to local industry, such as machines to help with the production of cotton or grain. Then, patents related to the railroad itself—couplers and, in the north, ways to clear snow from the tracks—started to appear. Following urbanization, newly connected places began patenting middle class consumption goods, such as medicine and furniture. This suggests that locally salient concerns, rather than more general problems on the technological frontier, drive patenting.

Together, these two chapters suggest transportation access has a positive effect on patenting because transportation forms a nexus that encourages local agglomerations, but the overall impact of larger transportation networks on innovation is unclear.

Changes to information flows also affect the behavior of both the electorate and politicians. Chapter 3 (with Steven Sprick Schuster)[4] moves from patenting to focus on the how mail delivery spread new information, studying the rollout of RFD in the early twentieth century. When deciding whether to vote and for whom to vote, coordinating with other voters, and interacting with their elected officials, potential voters rely on information from candidates, media sources, and peers.

Advancements in information distribution were especially important for rural residents, whose isolation acutely concerned policymakers.[5] This isolation was notably apparent in rural residents' lack of daily mail access. Since 1863, city dwellers enjoyed either at-home mail delivery or close proximity to post offices, while rural residents had to travel several miles to the nearest post office. These concerns led to the expansion of daily mail delivery to rural homes. Begun on an experimental basis in 1896, and rolled out across the country during the first decade of the twentieth century, RFD changed the flow of information to rural communities and the information networks within them.

Using a newly constructed panel data set, we find results consistent with the hypothesis that an increase in information to rural voters increases their political power. Though we rule out a large effect of RFD on voter turnout in Congressional elections, we find that RFD routes increased the competitiveness of Congressional elections and increased vote share for small parties. The results only occur in counties with local newspapers, suggesting that the main channel is a lowered cost to voters of acquiring political information. We also find that elected officials changed behavior in response to RFD allocation. The policy positions of congressional Representatives shifted toward stances associated with rural communities, which were primarily associated with Populist causes, including increased support for pro-temperance and anti-immigration policies.

My dissertation focuses on two ways expanding transportation and information networks connect rural areas to emerging national markets. By being more present in the political discourse, residents of rural areas are able to increase their ability to shape government policy, and, though not undoing advantages conferred by local clusters, increased connection also increases rural residents' participation in the national market for innovation.

<div align="right">Elisabeth Ruth Perlman, U.S. Census Bureau</div>

---

[4] Published by this Journal as Perlman and Sprick Schuster (2016).

[5] In his 1903 Annual Message to Congress, President Theodore Roosevelt said, "Rural free delivery, taken in connection with the telephone, the bicycle, and the trolley, accomplishes much toward lessening the isolation of farm life and making it brighter and more attractive."

REFERENCES

Agrawal, Ajay, and Avi Goldfarb. "Restructuring Research: Communication Costs and the Democratization of University Innovation." *American Economic Review* 98, no. 4 (2008): 1578–90.

Atack, Jeremy, Fred Bateman, Michael Haines, et al. "Did Railroads Induce or Follow Economic Growth?: Urbanization and Population Growth in the American Midwest, 1850–1860." *Social Science History* 34, no. 2 (2010): 171–97.

Atack, Jeremy, Michael Haines, and Robert A. Margo. "Railroads and the Rise of the Factory: Evidence for the United States, 1850–70." In *Economic Evolution and Revolution in Historical Time*, edited by Paul Rhode, Joshua Rosenbloom, and David F. Weiman, 162–79. Stanford: Stanford University Press, 2011.

Atack, Jeremy, Matthew Jaremski, and Peter L. Rousseau. "American Banking and the Transportation Revolution before the Civil War." *Journal of Economic History* 74, no. 4 (2014): 943–86.

Donaldson, Dave, and Richard Hornbeck. "Railroads and American Economic Growth: A 'Market Access' Approach." *Quarterly Journal of Economics* 131, no. 2 (2016): 799–858.

Fogel, Robert, W. *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore, MD: Johns Hopkins University Press, 1964.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. "Information Technology and the Distribution of Inventive Activity." In *The Changing Frontier: Rethinking Science and Innovation Policy*, edited by Adam Jaffe and Benjamin Jones, 169–96. Cambridge, MA: National Bureau of Economic Research, 2014.

Goldschlag, Nathan. "Investigations in the Study of Innovation and Firm Dynamics: Computational Simulations, Declining Entrepreneurship, and Natural Language Processing." Ph.D. thesis, George Mason University, 2015.

Gutberlet, Theresa. "Mechanization and the Spatial Distribution of Industries in the German Empire, 1875 to 1907." *Economic History Review* 67, no. 2 (2014): 463–91.

Harris, C. D. "The Market as a Factor in the Localization of Industry in the United States." *Annals of the Association of American Geographers* 44, no. 4 (1954): 315–48.

Howe, Daniel W. *What Hath God Wrought: The Transformation of America, 1815–1848*. Oxford History of the United States. Oxford: Oxford University Press, 2007.

Hudson, John C. *Plains Country Towns*. Minneapolis: University of Minnesota Press, 1985.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108, no. 3 (1993): 577–98.

Marshall, Alfred. *Principles of Economics*. New York: Macmillan and Co. 1890.

Perlman, Elisabeth R., and Steven Sprick Schuster. "Delivering the Vote: The Political Effect of Free Mail Delivery in Early Twentieth Century America." *Journal of Economic History* 76, no. 3 (2016): 769–802.

Sokoloff, Kenneth L. "Inventive Activity in Early Industrial America: Evidence From Patent Records, 1790–1846." *Journal of Economic History* 48, no. 4 (1988): 813–50.

The Dissertations of Shameel Ahmad, Réka Juhász, and Johannes Buggle: 2016 Alexander Gerschenkron Prize Competition

Fourteen new PhDs in economic history competed for the Gerschenkron Prize this year. Their research spans the globe, from Peru and Bolivia via Benin and Rhodesia to the Yangzi River. Many of them were born in the places that they study, and their innate understanding of local institutions and culture will push ahead the profession and improve all of our work.

Picking three among these dissertations was extremely difficult, and I found myself reminded of the judges at the Crystal Palace Fair. In 1851, industry experts, half of them British, examined exhibits from across the world and awarded prizes to what they thought were the most innovative. Their rankings, however, were shaped by their own distinct experiences, training, and tastes—as are mine. I therefore want to take this opportunity to welcome all of the 14 amazing new economic historians that have submitted their dissertations for the Prize this year. We are fortunate to have you and personally, I very much look forward to learning more from your research.

Now to the panelists. Thomas Henry Huxley a prominent defender of Darwin's theory of evolution, once explained that "The great tragedy of science—(is) the slaying of a beautiful hypothesis by an ugly fact" (Huxley 1870).[1] This is a key responsibility of our discipline. Grounded in both history and economics we establish facts, collect data, and combine them in rigorous tests to slay even the most beautiful hypotheses.

The three dissertations on this panel have done a lot of collecting facts, analyzing data, and some slaying. Here they come (in certified random order, according to Ray and Robson (2016), i.e., I rolled a dice).

SHAMEEL AHMAD

For his thesis on **"Demography and Economic Development in Colonial South Asia,"** Shameel Ahmad has collected an enormous amount of archival data to investigate the role of demographic changes in the development of late colonial India. In 1826, the Reverend Malthus described the mechanism that kept pre-industrial populations and material resources in equilibrium.

For Malthus, India and China were examples of the evil consequences of early and universal marriage: persistent poverty, famines, and epidemics. In that sense, India and China form a contrast for European exceptionalism, where decisions to delay marriage and have fewer children laid the foundations for economic growth.

Ahmad's dissertation presents new archival materials to challenge this view. New district-level data on births, deaths, and local monsoon rainfall for southern India between 1871 and 1931 show that fertility and death respond to rainfall shocks. Comparisons with early modern France and England, however, show that differences were much more nuanced than the traditional view suggests. South India was not an

---

[1] Born 4 May 1825, died 29 June 1895, grandfather of the author Aldous Huxley.

outlier in terms of its responsiveness to shocks, but instead was exposed to larger shocks due to its reliance on the monsoon.

These findings motivate a careful demographic exercise to correct existing estimates of vital rates over time and across space. Here, Ahmad accounts for under-counted population flows (in the form of births and deaths by age) with population stocks at the decadal census. By this process, Ahmad has created new annual district-level panel data for birth rates and age-specific mortality rates for more than 150 districts in the five largest provinces of British India. These data cover a staggering 200 million people between 1881 and 1931.

Ahmad also documents a dramatic response to the arrival of the railroad in India: Mortality fell steeply and became less responsive to extreme monsoon shocks, and fertility rose slightly. In the long-run, districts with railroads became 6 percent more densely populated. A final chapter introduces new data on fortnightly commodity prices during the 1870s. These data show that grain markets in British India transmitted prices much like those in eighteenth-century France, again calling into question European exceptionalism.

This is an impressive thesis, and if I had to push harder it would be on the links between demography and economic development. In its current state, Ahmad's work is very heavy on demography, probably because he (rightly) wants the rest of us to think more carefully about the dynamics of population growth, and push us to incorporate analytic tools of demography in our work. For example, Ahmad tells us that colonial India was in many ways comparable to European countries at a similar stage of development. This begs the question why India did not undergo its own demographic transition under colonial rule, and why living standards continued to stagnate.

This research agenda will have even greater impact when Ahmad strengthens its links with economic development.

RÉKA JUHÁSZ

Réka Juhász's dissertation on **"Temporary Protection, Technology Adoption and Economic Development"** addresses a first-order question in the economics of development and international trade: Can temporary protection encourage the development of infant industries?

Empirical evidence on this question is extremely scarce because governments typically grant trade protection in response to lobbying from the industries that benefit from such protection. This makes it nearly impossible to identify the effects of trade protection, so that existing literature has relied primarily on simulations (e.g., Baldwin and Krugman 1986; Irwin 2000).

Juhász addresses the challenge head on. Her thesis uses geographic variation in the effectiveness of Napoleon's blockade of Britain (1803–1815) as a source of exogenous variation to examine the effects of trade protection on mechanized cotton spinning in France. Napoleon's blockade was effective in Northern Europe, and imports from Britain fell by a factor of five. Southern Europe, however, was different, as Spain's insurgency against Napoleon weakened the effectiveness of the blockade.

Juhász uses this variation to investigate the effects of trade protection on mechanized cotton spinning, an industry where France lagged pitifully behind Britain at the

beginning of the Napoleonic wars. Yarn spun in France cost twice as much as yarn spun in Britain, and France was a net importer of cotton goods.

To perform the analysis, Juhász used hand-written archival sources to construct panel data on variation in production capacity of cotton spinning and weaving, wool spinning and weaving, and on the leather industry. To quantify changes in trade routes during the blockade, she extracted port level shipping data from *Lloyd's List*, one of the oldest newspapers in the world, and used text matching algorithms to track journeys between Britain and Continental Europe.

These data reveal a large increase in production capacity for French départéments (regions) that benefitted from additional protection. For example, moving from the 25th to the 75th percentile of the trade shock was associated with an increase in capacity that is comparable to average capacity by the end of the end of the blockade. Placebo tests for wool spinning and leather tanning, industries that were related to cotton spinning but less affected by trade shocks, show no comparable effects. Robustness checks examine the effects of factor prices (as in Allen 2009), and shocks in labor supply (measured by the number of conscripted men).

This work brings urgently needed empirical evidence to a policy debate with enormous welfare implications. It also demonstrates exceptional creativity, courage, and hard work. Two other chapters—on agglomeration externalities and on inter-industry linkages—are not quite there yet, but they also pursue difficult and important questions and a lot of intellectual honesty rigor. I have no doubt that Juhász will convert these two chapters into similarly outstanding papers.

JOHANNES BUGGLE

Johannes Buggle's dissertation **"Essays on Culture, Institutions and Long-Term Development"** is a worthy representative for the many dissertations this year that explore the historical origins of culture.

In the interest of time, I will focus on Chapter 1, which is particularly Gerschenkronian, and joint work with Steve Nafziger, the 2007 winner of the Gerschenkron Prize. In this chapter Buggle asks whether exposure to serfdom in the eighteenth and nineteenth centuries may be the cause of persistent under-development in Russia today.

To measure variation in the intensity of serfdom, the paper uses micro-level data for nearly 500 municipalities in 1861, immediately before the emancipation of peasants. The authors use Catherine the Great's 1764 Decree on Serfs as a source of exogenous variation in the intensity of exposure to serfdom. The Decree expropriated most monasteries and transferred almost 2 million serfs from private to state control, which, at the time, was a milder form of labor coercion. IV estimates use the pre-Decree locations of monasteries as an instrument for serfdom. These estimates show that locations that were more exposed to serfdom in the nineteenth century, are significantly poorer today, measured by household expenditures and wealth.

To help us understand the mechanisms of persistence, the paper also examines what happened in the 150-year gap between 1861 and today. For example, the paper shows that cities with higher historical levels of serfdom had fewer people and grew less throughout the twentieth century. Serf-cities also have higher levels of historical inequality and they are less likely to invest in public goods, primarily education.

Exposure to serfdom was also associated with lower degrees of industrial development throughout the Soviet period. These findings indicate that—contrary to the prevailing view—the roots of underdevelopment in the former Soviet Union extend far beyond Communism.

Three other essays in Buggle's dissertation are equally impressive. *Law and Social Capital* examines the long-run effects of exposure to the Napoleonic code civil on modern levels of trust in Germany. Buggle shows that German regions that were exposed to the code for a longer period of time have higher levels of social trust today, as measured by survey data in the Socio-Economic Panel. Interestingly, there is no clear effects of the code on economic development, or inequality. Exposure to the code was, however, negatively correlated with political fraud in elections between 1871 to 1900, which suggests that the code affected social norms if not economic development.

Chapter 3 on *Irrigation and the Origins of Collectivism*, examines the link between ancient modes of economic production and the formation of cultural beliefs, in the spirit of Robert Boyd and Peter J. Richerson (2004). Chapter 4 (co-authored with Ruben Durante) studies the potential effects of climatic variability between 1500 and 1750 on social trust and on the quality of political institutions today. A particular strength of that paper is that it examines the interactions between culture and institutions.

This dissertation includes four strong essays, which together form an outstanding body of work. Yet, some of the essays need deeper analyses of historical mechanisms of persistence and cultural transmission, as well as alternative mechanisms. The thesis already takes steps in this direction, and finishing the job will require little more than a final push.

Let me conclude by congratulating the finalists on a job well-done and please join me in welcoming them and all other new PhDs to our profession.

PETRA MOSER, *New York University*

REFERENCES

Allen, Robert C. *The British Industrial Revolution in Global Perspective*. Cambridge: Cambridge University Press, 2009.

Boyd, Robert, and Peter J. Richerson. *The Origin and Evolution of Cultures*. New York: Oxford University Press, 2004.

Baldwin, Richard E., and Paul Krugman. "Market Access and International Competition: A Simulation Study of 16K Random Access Memories." NBER Working Paper No. 1936, Cambridge, MA, 1986.

Huxley, Thomas Henry. Presidential Address at the British Association, "Biogenesis and abiogenesis" (1870) in Collected Essays, 1890, vol. 8, p. 229.

Irwin, Douglas A. "Did Late-Nineteenth-Century U.S. Tariffs Promote Infant Industries? Evidence from the Tinplate Industry." *Journal of Economic History* 60, no. 2 (2000): 335–60.

Ray, Debraj, and Arthur Robson. "Certified Random: A New Order for Co-Authorship." NBER Working Paper No. 22602, Cambridge, MA, 2016.

Demography and Economic Development in Colonial South Asia

For the classical economists, demography was fundamental to the theory and measurement of economic growth. Perhaps the most celebrated of these is the Reverend Thomas Robert Malthus, who described the mechanisms that kept pre-industrial populations and material resources in equilibrium. Part of his analysis, like that of his intellectual heirs today, rests on the contrast between Western Europe and places like China and India. The latter societies, with early and universal marriage and hence a weak preventive check, were doomed to adjustments through the positive checks of famine, epidemic, and war. But how appropriate is the Malthusian model for understanding the world outside of Europe? The evidentiary basis for Malthusian mechanisms in India in particular is thin, and yet, if we are to take Malthus seriously, the implications for understanding Indian growth are profound.

My primary contributions in this dissertation are providing evidence on Malthusian population dynamics in colonial India (Chapter 1), developing a new dataset of demographic statistics from flawed archival sources (Chapter 2), and testing the responsiveness of Indian population change to the gains from trade associated with the British-built railroads (Chapter 3). I briefly summarize a related study on agricultural market efficiency in late nineteenth-century India (Chapter 4).

In the first chapter, I study the relationship between rainfall shocks and birth and death rates in the Madras Presidency of southern India over 1871–1931. In the basic Malthusian model, population levels are endogenous to economic productivity because vital rates respond to real wages. Testing the implications of this model over the long run is difficult, given the endogeneity of population flows and stocks to prior economic and demographic change: economic historians have therefore tended to focus on the short-run predictions of the model, namely sign restrictions on the elasticities of fertility (positive) and mortality (negative) to good economic shocks.

I employ newly digitized data on births and deaths collected by colonial public-health administrators beginning in the 1870s. These provide a unique view, at a temporally- and spatially-disaggregated level, of vital events for this period of Indian history.[1] Indeed, it is exceptionally rare among today's developing countries to be able to trace vital rates backwards earlier than 1950.Demographic historians such as Tim Dyson (2002) have been critical of this source, primarily for its unevenness in reporting: the incomplete recording of vital events (under-registration) varied in ways that vitiate cross-sectional and cross-time comparisons. I side-step these concerns by computing short-run deviations from moving averages, regressing them in a distributed-lag framework on measures of district-specific rainfall shocks.

Shameel T. Ahmad, Florence Levy Kay Fellow, Department of Economics, Brandeis University, Waltham, MA. E-mail: shameel@brandeis.edu. I would like to thank my advisors Timothy Guinnane, Naomi Lamoreaux, and Daniel Keniston for their unstinting support while writing this dissertation. In addition, I gratefully acknowledge financial support from the Economic History Association, the National Science Foundation (Grant #SES-1426800), Yale's Economic Growth Center and the Cowles Foundation.

[1] In fact, very few of today's developing countries have such statistical records before the middle of the twentieth century, let alone for the nineteenth century.

I find that mortality in Madras does indeed respond strongly to local rainfall shocks in a manner consistent with the Malthusian model, particularly to the northeast monsoon rainfall that is the critical input for rice production in southern India. Surprisingly, I also evidence of a small, but statistically significant, fertility response: given the prevailing view of the Indian demographic system as one featuring uncontrolled fertility, this suggests some scope for the operation of a preventive check along the lines held to be unique to Western Europe, even if it was comparatively weaker. Conditional on the size of the underlying shock, southern Indian vital rates are not dramatically more responsive than those for seventeenth and eighteenth century England and France, estimated by David Weir (1984).[2] Two important features of European development in the nineteenth century were the progressive diminution of both the underlying shocks, as markets become increasingly integrated and efficient, and the demographic elasticities, as fertility, nuptiality, and mortality become increasingly decoupled from agricultural shocks. However, Madras does stand out for the magnitude and volatility of the shocks from year to year. I conclude from this analysis that the basic Malthusian mechanisms deserve attention in understanding South Asia's economic history; moreover, I argue that the underlying vital registration data deserve closer attention in order to be rendered useful for statistical analysis.

In my second chapter, I develop a new technique to recover unbiased estimates of district-level vital rates for British India from under-registered vital event counts and decennial Census enumerations of district populations by age. I study around 150 districts in five major provinces between 1881 and 1931, where a complete time-series of vital events can be constructed from statistical reports, with suitable adjustments for areal changes: this covers a population of more than 200 million. The goal is to generate a dataset of district-level birth and death events that is internally coherent—that is, one that obeys simple accounting relationships that must hold between stocks (populations) and flows (vital events).

Critically for my purposes here, annual counts of deaths by district are reported not only in aggregate but also in relatively narrow age intervals and by gender. Assuming that under-registration does not jump from year to year, mapping from the age distribution of adult males and females observed in a district at two successive Censuses (say, 1891 and 1901) is a matter of accounting for age-specific mortality experienced by each single-year cohort in each intervening years and the fact that survivors must be ten years older.[3]

Similarly, the number of children under the age of 10 in 1901 is a function of birth rates, infant- and child-mortality rates in the preceding decade. Broadly, my approach is to simulate a coherent population system with arbitrary annual flows of births and deaths, such that the simulation matches archival data on stocks (age-enumerated Census populations) and flows (annual births and age-enumerated deaths) when under-registration,

---

[2] By the nineteenth century, however, European economic demography began to change more dramatically, featuring the progressive diminution of both the underlying shocks, as markets become increasingly integrated and efficient, and the demographic elasticities, as fertility and mortality become increasingly decoupled from agricultural shocks.

[3] Strictly speaking, this mapping requires the assumption that populations are closed to migration. I argue that this a good first approximation for my sample districts and period, based on statistics that enumerate district populations by district-of-birth.

age mis-statement and the empirical regularities of the age-mortality relationship are accounted for.

I expand on these three substantive assumptions, as they are crucial to understanding the nature of the exercise. The first is that registration (for a district-gender-age cell) is constant within a decade. I allow in principle for differential birth and death registration by district, gender and age, and find that these do differ dramatically. For example, I find that the best-performing districts miss only a quarter of vital events, whereas the worst miss more than half. Female births are consistently under-reported in parts of north India, and under-registration appears to decline monotonically with age (with half of infant deaths going under-counted, compared with a fifth of elderly deaths). Second, I assume that age mis-statement patterns do not vary from one Census to the next within a district, inverting the simulated population age distribution to match that reported in age-enumerated data.[4] Finally, I assume that the age-mortality relationship for any district-gender-year cell is well-described by a suitably flexible parametric relationship: I find that newly-developed two-parameter model life tables from John Wilmoth et al. (2012) fit remarkably well, imposing sufficient structure on the within-year age-mortality relationship that I can allow overall mortality to vary arbitrarily across years, consistent with Indian districts being subject to idiosyncratic crisis mortality.

Practically speaking, a high-performance computing approach is required, as the problem is high dimensional and non-convex. I therefore employ simulated annealing, a stochastic optimization procedure, in parallel across districts and decades. I find that overall about two-thirds of deaths and 60 percent of births were registered in India, with wide variation across time and districts in the levels and trends of birth and age-specific death rates. I compare my demographic estimates at an aggregate level to those from the literature, which have not been built up from the vital registration data, to demonstrate their essential consistency. Finally, I discuss how these data can be used to answer old questions in Indian history, for example on the geography, timing, and causes of India's great mortality decline in the early twentieth century.

In my third chapter, I use these corrected demographic indicators to examine how the arrival of the railroads across India affected vital rates and population growth. The British-built railroads are celebrated as one of the greatest benefits of imperial rule, and yet historians often emphasize their role in agricultural immiseration and overall anemic rates of economic growth. I argue that Malthusian demographic responses may be one way of reconciling per-capita income stagnation with major structural change.

Trade theory delivers straightforward predictions about the consequences for falling transport costs on local incomes: as factors reallocate to sectors of comparative advantage and gains from specialization and exchange are realized, real incomes rise. Indeed, this is the insight Dave Donaldson (2010) brings to studying the impact of the colonial railroads, demonstrating that real revenue per acre rose as the network expanded. Indeed, this is the insight Donaldson (2010) brings to the study of the British railways in this very context, demonstrating convincingly that real income, measured as deflated gross agricultural product per cropped acre. However, the logic of Malthus suggests that gains from trade will take a different form where fertility and mortality respond to income: they will be reflected in the short-run in changes in vital rates, and in the

---

[4] Indeed, the patterns I find are not qualitatively dissimilar from those discussed by demographers studying modern India. See, for example, Bhat (1990) on age misstatement in India's 1970 census.

long run only in population levels.[5] Formally, population levels may also change due to differential net migration. However, Census enumerations of the birthplace of district residents suggests that the railroad played little role in encouraging permanent cross-district migration, at least in my sample districts and period. I demonstrate this formally using a benchmark neo-Ricardian model, of the type developed by Jonathan Eaton and Samuel Kortum (2002), augmented with a Malthusian dimension in which population is endogenous to the real wage.

I make three predictions using this model and test them in my new dataset by exploiting the heterogeneity in the timing of the arrival of railroads across Indian districts. First, the combined logic of the Ricardian and Malthusian models suggest that the railroads should be followed by rising fertility and lower mortality. This is indeed what I find, with a 6 percent increase in the birth-rate and an 8–10 percent decrease in the death rate. Second, the long-run effect of the railroads should be reflected in denser populations: I conclude that districts are about 6 percent denser post-railroad than they would have been in the absence of the railroad, a large effect relative to overall population growth over this period of about 40 percent. Third, I test the hypothesis that railroads reduce vital-rate sensitivity to bad rainfall shocks: the railroad entirely eliminates the responsiveness of mortality to negative monsoon shocks, though I find no comparable impact for fertility. Finally, I use the model to conduct a back-of-the-envelope calculation: how much higher might Indian wages have been if the gains from trade had not been absorbed by expanding populations? My analysis suggests that counterfactual real wage gains may have been about 3 percent: slightly larger than independently-measured estimates of real wage growth over the same period. I conclude that the demographic margin of adjustment to market integration is substantial, and can reconcile why the railroad's income effects appeared so small even while their economic effects were substantial.

Finally, in the fourth chapter, I tackle a narrower question: How efficient were agricultural commodity markets in India in the 1870s, and how did they respond during the Great Famine that falls towards the end of that decade? A minor cottage industry has sprung in economic history engaged in the comparative study of market integration and efficiency across countries and time. For India in particular, the primary data source used to estimate the degree of market efficiency to date has been annual price data, which I argue is necessarily too coarse to pick up adjustments in prices that take place due to sub-annual arbitrage. Using a newly digitized set of fortnightly prices for rice and wheat for about 70 Indian cities between 1871 and 1883 and an error-correction model, I study the synchronicity and bilateral adjustment of rice and wheat prices across all bilateral pairs. I estimate adjustment parameters that are comparable in magnitude to those in the literature from France in the seventeenth century and Finland in the nineteenth century. In examining how these parameters change during the Great Famine of 1876–1878, I find that markets moved more quickly to re-converge to spatial equilibrium than they did during famines in either of the aforementioned historical contexts. I conclude from this study that markets in India at the dawn of its railway age operated at similar levels of efficiency to those of European countries at similar levels of

---

[5] The railroads can also drive local population change if they affect net migration. However, my robustness regressions suggest that these effects are statistically insignificant and small in magnitude relative to changes in vital rates.

development, and that they seem to have operated to relieve distress during one of the largest famines in its history.

Beyond developing methods for working with flawed demographic statistics and bringing new arguments and evidence to debates about Indian history, my dissertation holds a more general lesson for economic historians. The nature of endogenous demographic adjustment to structural economic change in pre-transitional societies implies that population may be central, as a matter of theory and measurement, to studying long-run economic development for most of the world until quite recently.

<div align="right">

Shameel T. Ahmad, *Brandeis University*

</div>

REFERENCES

Bhat, P. N. Mari. "Estimating Transition Probabilities of Age Misstatement." *Demography* 27, no. 1 (1990): 149–63.

Donaldson, Dave. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." NBER Working Paper No. 16487, Cambridge, MA, 2010.

Dyson, Tim. "Famine in Berar, 1896–7 and 1899–1900: Echoes and Chain Reactions." In *Famine Demography: Perspectives from the Past and Present,* edited by Cormac Ó Gráda and Tim Dyson, 93–112. Oxford: Oxford University Press, 2002.

Eaton, Jonathan, and Samuel Kortum. "Technology, Geography, and Trade." *Econometrica* 70, no. 5 (2002): 1741–79.

Weir, David R. "Life Under Pressure: France and England, 1670–1870." *Journal of Economic History* 44, no. 1 (1984): 27–47.

Wilmoth, John, Sarah Zureick, Vladimir Canudas-Romo, et al. "A Flexible Two-Dimensional Mortality Model for Use in Indirect Estimation." *Population Studies* 66, no. 1 (2012): 1–28.

Temporary Protection, Technology Adoption, and Economic Development: Data and Evidence from the Age of Revolution in France

A long-standing debate in economics is centered on the question of whether certain industries can benefit from temporary trade protection in industrial follower countries. The idea, widely known as the infant industry argument, has a long tradition in the history of economic thought.[1] However, assessing the empirical relevance of the infant industry argument has proven difficult for two reasons. First, as policy is generally implemented at the country-wide level, even if the industry becomes competitive in the long run, it is difficult to answer the question of whether the industry would have become competitive anyway. Second, in the case of a specific policy intervention, it

Réka Juhász, Assistant Professor, Economics, Columbia University, New York, NY. E-mail: rj2446@columbia.edu. This dissertation was completed under the supervision of Silvana Tenreyro and Steve Pischke at the London School of Economics.

[1] Early proponents include U.S. founding father Alexander Hamilton (1791), and the nineteenth-century economist, Friedrich List (1841).

is not possible to disentangle the effect of the economic mechanism at work from the efficacy of implementation.[2]

This dissertation examines the effect of temporary trade protection during the Napoleonic Wars (1803–1815) on industrial development in France in the nineteenth century. I study how protection affected the development of mechanized cotton spinning, the first industry to adopt modern, factory-based production technology in France. I then turn to examining how the location of this first modern industry affected the spatial distribution of other, technologically proximate sectors which mechanized decades later.

To answer these questions, I digitized data on a number of industries, observed at various points in time between the late eighteenth and mid-nineteenth century. Sectors covered include cotton spinning and weaving, wool spinning and weaving, and leather tanning. The dataset is constructed from large-scale, industrial surveys conducted between 1790–1815. Data from the Napoleonic period are widely acknowledged to be of very high quality; Jean-Claude Perrot (1975) refers to them as the "Golden Age of French Statistics," while George Grantham (1997) argues that the surveys from this period are of a quality as high as any survey on economic activity conducted within the next century. I link these surveys to two industrial censuses which were conducted in the 1840s and 1860s.[3] For each industry, I observe employment at the level of the department, along with a number of other variables specific to the particular sector. For mechanized cotton spinning, I observe the number of mechanized spindles—the relevant measure of physical capital in the industry.

Textiles was one of the leading industries during the Industrial Revolution. Its patent intensity in Britain was second only to engines (Nuvolari and Tartari 2011). In Britain, cotton textiles manufacturing is estimated to have accounted for 25 percent of Total Factor Productivity (TFP) growth between 1760–1860 (Crafts 1985, p. 131). In France, textile manufacturing accounted for 60 percent of all manufacturing employment in 1840.[4] In contrast to the changes taking place in textiles, leather was one of the industries in which change seems to have been almost imperceptible. It was one of the least patent intensive industries, and it mostly remained rurally organized throughout my time period of interest (Bonin and Langlois 1997).

A number of features of early nineteenth century France makes the setting an appealing one in which to test infant industry mechanisms. First, mechanized cotton spinning in early nineteenth century France was an infant industry. It was technologically backward relative to Britain and French mechanized spinners were producing cotton yarn at a significantly higher price than their counterparts in Britain.[5] A number of key technological breakthroughs in Britain had mechanized the production of cotton spinning in the second half of the eighteenth century (Allen 2009). Despite the fact that

---

[2] The literature to date generally uses estimated model parameters to simulate the counterfactual of no-protection. Examples include Baldwin and Krugman (1986, 1988), Head (1994), Luzio and Greenstein (1995), Hansen, Jensen, and Madsen (2003), and Irwin (2000). This approach implicitly evaluates both the underlying economic mechanism—in general learning-by-doing production externalities—together with the efficacy of policy.

[3] These data were digitized by Chanut et al. (2000).

[4] Manufacturing employment calculated using data from Chanut et al. (2000).

[5] Data comparing the price of cotton yarn in Britain and France from *Archives Nationales*, Series F12/533.

French entrepreneurs had widespread knowledge of and access to the new technology being developed across the Channel, it was not adopted on a wide scale prior to the Napoleonic Wars (Aspin and Chapman 1964). One explanation for this could be infant industry mechanisms, which do not make it profitable for any single entrepreneur to enter the market at current prices, even though the industry could become competitive if all other entrepreneurs entered simultaneously.

Crucially, mechanized cotton spinning was the first industry in which modern factory-based production methods were employed (Allen 2009). This aspect of the historical context is important as any infant industry model relies on some form of increasing returns to scale.[6] While it is unlikely that these types of agglomeration economies will be present in traditional, "cottage industry" manufacturing, there is a large body of historical evidence that documents various forms of learning by doing externalities for mechanized cotton spinning (Mokyr 2009).

Second, exogenous, within country variation in trade protection makes it possible to compare outcomes in areas in which the cost of trading with the industrial leader, Britain, increased to a larger extent, to areas in which the increase was smaller. The fact that trade protection was not driven directly by the actions of the policymaker makes it possible to disentangle the economic from the political mechanism.

The unique way in which blockade against British goods was implemented by Napoleon resulted in a differential shock to the costs of trading with the industrial leader across different regions of the French Empire. In particular, Napoleon did not have the naval military power to impose a standard blockade of the British Isles whereby the French navy would surround British ports. He instead used his direct or indirect influence over much of Continental Europe to attempt to stop British goods from entering the mainland. Ports were closed to British trade, and the military was active in enforcing the blockade along the coastline. In practice, holes in the system opened up almost immediately, and instead of achieving the original goal of stopping trade flows between Britain and the Continent, the blockade displaced trade to more circuitous, and hence more expensive routes.

The key to variation in trade protection across the French Empire is the uneven geographic success of the blockade across Europe. The blockade was generally effective in Northern Europe with exports from Britain falling fivefold from peak to trough (Crouzet 1987). Trade intended for Northern European markets however, was diverted to Southern Europe. British exports to the region increased threefold, as Napoleon's inherent military weakness and the idiosyncratic political event of the Spanish insurgency against Napoleonic rule meant that the blockade against British trade was unsuccessful. The geographic asymmetry in the success of the blockade outside of France meant that trade flows to different parts of the French Empire were disrupted to a different extent.

In the north of France, effective distance between a given region and London increased markedly, as trade was diverted either to unreliable indirect routes through German regions, or through Southern Europe. In the southern regions of France, effective distance to London changed to a far smaller extent, as trade routes stayed more or less the same. Changing trade protection from Britain during the Napoleonic Wars was

---

[6] For example, Krugman (1987), Lucas (1988), Matsuyama (1992), and Young (1991) analyze models in which source of increasing returns to scale is a learning-by-doing production externality.

thus driven by one of the best-documented empirical regularities in economics; the fact that trade diminishes dramatically with distance.

To answer the question of how protection from trade with Britain affected the development of mechanized cotton spinning in France, the first part of the empirical strategy employs a difference in difference (DD) estimator with continuous treatment intensity. The strategy compares the size of mechanized cotton spinning capacity across regions which were exposed to smaller or larger increases in the cost of trading with Britain (trade-cost shock for short), before and after the Napoleonic Wars. I find that areas which received a larger trade cost shock during the Napoleonic Wars increased production capacity in mechanized cotton spinning to a larger extent than areas which received a smaller shock. The estimated effect is large and statistically significant.

If an infant industry type mechanism was inhibiting French entrepreneurs from switching into the new technology, we would expect not only that departments scaled up in mechanized spinning capacity during the period of temporary protection, but also that they remained profitable in the long-run, once pre-blockade variation in trade protection was restored. To examine this question, in the second part of the empirical analysis, I ask to what extent temporary trade protection had long-term effects on the industry. I examine persistence in the location of the industry within France, and ask whether regions with higher spinning capacity after the blockade were more productive 30 years later.

To do so, I exploit the fact that the post-blockade location of the cotton industry was determined to a large extent by the historical accident of the Napoleonic Wars. Using the trade cost shock as an instrument for the post-blockade location of the cotton industry, I find that regions with a higher level of spinning capacity at the end of the Napoleonic Wars, also had higher capacity 30 years later. The estimated effect is large and significant. Furthermore, areas with higher post-blockade production capacity had more productive firms 30 years later. Again, the effect is large and significant.

In the second part of the dissertation, I turn to examining the importance of inter-industry linkages in nineteenth century France. Recent years have seen a renewed interest in the idea that specialization patterns across countries display some level of indeterminacy. In an influential paper, Ricardo Hausmann, Jason Hwang, and Dani Rodrik (2007) have shown that countries that have a revealed comparative advantage in goods which closely resemble the mix of goods that rich countries tend to export grow faster than countries which are similar in terms of observables, but happen to produce more typical "poor country goods." The authors claim that this finding, to the extent that it is not driven by unobservable fundamentals, is consistent with a world in which indeterminacy in specialization is empirically important.

Theoretically, a variety of models deliver predictions in which fundamentals do not uniquely pin down specialization patterns across space.[7] In the case of inter-industry knowledge spillovers, for example, specializing in a given product will, all else equal, increase productivity in technologically similar products. In this way, a region which specializes in a specific industry will become a more attractive location for other

---

[7] The indeterminacy can be driven by production externalities as in Lucas (1988), Matsuyama (1992), Krugman (1987), and Young (1991), coordination failures as in Rodrik (1996) or Rodríguez-Clare (1996), or new-economic geography mechanisms discussed in Fujita, Krugman, and Venables (2001).

industries which are technologically similar. The scope for large effects on specialization crucially depends on the existence of strong inter-industry linkages. These linkages are generally hard to estimate, as a host of confounding factors makes disentangling the different forces at work notoriously difficult.

To answer the question of whether the location of mechanized cotton spinning affected the spatial distribution of other, technologically proximate industries which mechanized decades later, I exploit exogenous variation in the location of the first modern sector in France, mechanized cotton spinning, to assess the strength and importance of inter-industry linkages. I examine the co-location of firms which produced yarn using different types of fibers (wool, flax, and silk).

Mechanization of production methods in these spinning sectors lagged developments in cotton spinning by decades for reasons inherent to the nature of the different fibers. Mechanizing downstream weaving took even longer. This aspect of the setting gives credibility to my identifying assumption, which requires all effects of the trade cost shock relevant to mechanization in textiles to be channeled through mechanized cotton spinning. I find that spinning sectors which were technologically more proximate to cotton spinning (flax and wool) co-located with cotton, while the same does not hold for sectors which were technologically more distant (silk). The point estimates are stable across time for a given industry. The estimated effect on silk is negative, though it is generally not significant at conventional levels.

In conclusion, this dissertation has found that temporary protection had a large and positive effect on the long-term location of the mechanized cotton spinning industry and technologically proximate sectors. What do the findings from this episode teach us about the effect of trade protection on industrial development of infant industries more generally? An appealing aspect of this episode is the extent to which it is general to the setting of structural transformation. For many of today's developed countries, unskilled-labor intensive textile manufacturing was the first industry into which labor flowed from agriculture at the early stages of industrialization.

However, this does not imply that trade policy will necessarily be an effective or desirable tool when implemented by a policymaker targeting specific sectors. Aside from political economy problems related to implementing infant industry promotion and the question of whether such a policy is welfare-enhancing, there is a further question pertaining to the extent to which the cotton spinning sector in early nineteenth century France may have been special. One reason why temporary protection may have had a large effect in France is that the cotton spinning industry had well-established access to raw inputs and also had a domestic downstream industry. The extent to which these results would hold in settings where industries are more isolated from their upstream suppliers and downstream markets is an interesting question for future research.

RÉKA JUHÁSZ, *Columbia University*

REFERENCES

Allen, Robert C. *The British Industrial Revolution in Global Perspective*. Cambridge: Cambridge University Press, 2009.

Aspin, C., and Chapman, S. D. *James Hargreaves and the Spinning Jenny*. Preston: Helmshore Local History Society, 1964.

Baldwin, Richard, and Paul R. Krugman. "Market Access and International Competition: A Simulation Study of 16K Random Access Memories." NBER Working Paper No. 1936, Cambridge, MA, 1986.

———. "Industrial Policy and International Competition in Wide-bodied Jet Aircraft." In *Trade Policy Issues and Empirical Analysis*, edited by Robert E. Baldwin, 45–78. Chicago: University of Chicago Press, 1988.

Bonin, Serge, and Langlois, Claude. *Atlas de la Révolution française: Economie* (Vol. 10). Paris: Editions de l'Ecole des hautes études en sciences sociales, 1997.

Chanut, J. M., J. Heffer, J. Mairesse, et al. L'industrie française au milieu du 19 e siècle (les enquêtes de la statistique générale en France). Paris: Recherches d'histoire et de sciences sociales, 2000.

Crafts, N. F. R. *British Economic Growth During the Industrial Revolution*. Oxford: Clarendon Press, 1985.

Crouzet, François. L'économie Britannique et le Blocus Continental: 1806–1813. Paris: Presses universitaires de France, 1987.

Fujita, Masahisa, Paul R. Krugman, and Anthony J. Venables. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press, 2001.

Grantham, George. "The French Cliometric Revolution: A Survey of Cliometric Contributions to French Economic History." *European Review of Economic History* 1, no. 3 (1997): 353–405.

Hamilton, Alexander. *Report on Manufactures*. Report to the House of Representatives, 1791.

Hansen, Jørgen D., Camilla Jensen, and Erik S. Madsen. "The Establishment of the Danish Windmill industry—Was it Worthwhile?" *Review of World Economics* 139, no. 2 (2003): 324–47.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. "What You Export Matters." *Journal of Economic Growth* 12, no. 1 (2007): 1–25.

Head, Keith. "Infant Industry Protection in the Steel Rail Industry." *Journal of International Economics* 37, no. 3 (1994): 141–65.

Irwin, Douglas, A. "Did Late-Nineteenth-Century US Tariffs Promote Infant Industries? Evidence from the Tinplate Industry." *Journal of Economic History* 60, no. 2 (2000): 335–60.

Krugman, Paul. "The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies." *Journal of Development Economics* 27, no. 1 (1987): 41–55.

List, Friedrich. *The National System of Political Economy*. London: Frank Cass, 1841.

Lucas, Robert E., Jr. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1988): 3–42.

Luzio, Eduardo, and Shane Greenstein. "Measuring the Performance of a Protected Infant Industry: The Case of Brazilian Microcomputers." *Review of Economics and Statistics* 77, no. 4 (1995): 622–33.

Matsuyama, Kiminori. "Agricultural Productivity, Comparative Advantage, and Economic Growth." *Journal of Economic Theory* 58, no. 2 (1992): 317–34.

Mokyr, Joel. *The Enlightened Economy: An Economic History of Britain 1700–1850*. New Haven, CT: Yale University Press, 2009.

Nuvolari, Alessandro, and Valentina Tartari. "Bennet Woodcroft and the Value of English Patents, 1617–1841." *Explorations in Economic History* 48, no. 2 (2011): 97–115.

Perrot, Jean-Claude. L'âge d'Or de la Statistique Régionale Française (An IV-1804). Paris, France: Société des Études Robespierristes, 1975.

Rodríguez-Clare, Andrés. "The Division of Labor and Economic Development." *Journal of Development Economics* 49, no. 1 (1996): 3–32.

Rodrik, Dani. "Coordination Failures and Government Policy: A Model with Applications to East Asia and Eastern Europe." *Journal of International Economics* 40, no. 1 (1996): 1–22.

Young, Alwyn. "Learning By Doing and the Dynamic Effects of International Trade." *Quarterly Journal of Economics* 106, no. 2 (1991): 369–405.

Essays on Culture, Institutions and Long-Term Development

This dissertation is a collection of four essays that contribute to the literature on the historical origins of economic development and its determinants. It studies the long-term persistence of four diverse historical events, thereby addressing three main research questions: (1) How do past institutions affect long-term economic development? (2) How do formal institutions interact with cultural norms? (3) What are the historical origins of cultural norms, in particular social trust and collectivism, and how do these norms persist?

Chapter 1 analyses the long-term economic consequences of Russian serfdom and provides evidence of the adverse effects of labor coercion in a novel context. Chapter 2 uses the "historical experiment" of the introduction of the Code Napoleon in Germany to examine whether formal institutions that support cooperation affect norms of trust. Chapter 3 links the evolution of collectivist beliefs to the adoption of irrigation in pre-industrial agriculture. Chapter 4 shows that pre-industrial climate fluctuations shaped formal and informal means of social cooperation within Europe.

Besides focusing on persistence, all chapters are also linked by a common methodology. The articles make use of micro-data, both historical and contemporary, that has partly been collected and constructed for the purpose of this thesis. Moreover, throughout this dissertation I apply the methods of micro-econometrics, including historical experiments, regression discontinuity designs, and instrumental variables, which help to identify causal relationships between the past and the present.

I will briefly summarize each chapter in the following.

LONG-RUN CONSEQUENCES OF LABOR COERCION: EVIDENCE FROM RUSSIAN SERFDOM

The first chapter of this dissertation—co-written with Steven Nafziger (Williams College)—asks whether labor coercion in the Russian Empire has left a legacy on economic development in Eastern Europe. While many scholars have argued that

Johannes C. Buggle, Postdoctoral Researcher, University of Lausanne, Switzerland. E-mail: johannes.buggle@unil.ch. This dissertation was completed under the supervision of Yann Algan at Sciences Po, Paris.

serfdom slowed down economic growth in the Russian Empire (see, e.g., Gerschenkron 1966), empirical evidence of short run, or even longer lasting economic effects is limited. We base our analysis of the economic consequences of serfdom on unique data that documents the intensity of labor coercion in almost 500 municipalities of the Russian Empire in 1858—just three years before the emancipation of peasants. We link the historical data to a rich set of contemporary and intermediate indicators of socio-economic development. For example, we make use of a geo-referenced household survey that measures economic wellbeing in 2006 and 2010, draw on population data of Russian cities in different years from 1929 to 1990, and employ data on public goods, economic inequality, and urbanization and structural change around the turn of the nineteenth century.

Our empirical results from ordinary least squares regressions indicate that municipalities in which a larger share of the population were serfs in 1858 are poorer today, measured by contemporary household expenditure per capita and household wealth. These regressions take into account a rich set of observable district characteristics, including the suitability of the environment for agricultural production, measures of early economic development, as well as provincial fixed effects. To identify causality, we use variation in the intensity of serfdom resulting from the Decree on Serfs issued by Catherine the Great in 1764. This decree essentially expropriated most monasteries and transferred around 2 million serfs from private to state control, easing many of the institutional constraints that private serfs experienced. Using the number of monasteries that were closed as an instrumental variable, we find even larger effects of serfdom on contemporary development.

Moreover, we document persistence of economic disparities between areas with high and low levels of serfdom throughout the twentieth century: cities in areas in which serfdom was more widespread were on average significantly smaller and did not catch up. Additional findings indicate that serfdom impeded structural change towards industrial production and urbanization. These agglomeration effects persisted during the Soviet period, potentially reinforced by Soviet policies that benefitted already urbanized areas. To sum, contrary to the prevailing explanation for the low levels of economic development in countries of the former Soviet Union that focuses on the adverse effects of Communism, our results suggest that parts of the underdevelopment originate from more distant institutional arrangements.

LAW AND SOCIAL CAPITAL: EVIDENCE FROM THE CODE NAPOLEON IN GERMANY[1]

Do people trust each other more if they face legal institutions that foster cooperation? Moreover, can such institutions have a long-term impact on a society's culture? In theory, legal institutions that enforce impersonal interactions reduce cheating behavior and increase cooperation and interpersonal trust (Tabellini 2008). Studying the relationship between institutions and norms of trust empirically is difficult, however, because institutions are usually endogenous to cultural norms. Chapter 2 analyses the long-term effect of the law on trust, using exogenous variation in the quality of legal institutions generated by the introduction of the Code Napoleon in parts of early nineteenth-century

[1] This chapter was published as Buggle (2016).

Germany. The Code Napoleon was created to spread the ideas of the French Revolution and to modernize the pre-existing social order of European societies. Its most revolutionary principle was to guarantee equality before the law, which helped to remove existing barriers to inter- and intra-class cooperation.

In Chapter 2 I show that individuals surveyed in the German Socio-Economic Panel have on average higher levels of social trust if they live in regions where the Code Civil was applied longer during the nineteenth century. However, this correlation could reflect pre-existing differences between areas that used the Code and those that did not. To address such concerns, I refine the empirical strategy and compare only respondents that live in neighboring communities that differed largely with respect to how long the French legal system was used, but were similar on many other dimensions. Regressions that use observations around this discontinuity confirm the positive impact of the Code Civil on contemporary trust. A falsification test that moves the discontinuity artificially, and horse races that compare the effect of the Code Civil to other Napoleonic institutional reforms, give additional support that the Code Civil influenced cooperation. Chapter 2 also explores several intermediate mechanisms. On the one hand, I find only limited evidence that the Code affected historical and contemporary economic development, or economic inequality. On the other hand, my empirical results suggest that the impartial legal institution eased the spread of norms of generalized morality: The Code Civil is negatively associated with political fraud in elections from 1871 to 1900, and goes along with faster evolving social associations in the second half of the nineteenth century.

IRRIGATION AND THE ORIGINS OF COLLECTIVISM

Previous research in economics has shown that individualist societies are more innovative, richer, and have better institutions (Gorodnichenko and Roland 2011). But why are some societies more individualistic, while others are more collectivistic? A prominent hypothesis in cross-cultural research argues that societies adopted those cultural norms that were most beneficial in pre-industrial agricultural economies (Boyd and Richerson 2004). In particular, in societies in which subsistence farming relied on group effort and conformity, collectivist cultural norms developed which reflect the relative importance of the group over the individual (Triandis and Gelfand 2012). A prime example of an agricultural technique that required group cooperation and conformity is the use of irrigation.

Chapter 3 tests whether descendants of societies that traditionally used irrigation have stronger collectivist beliefs today. I use ethnographic data to measure how important the use of irrigation in agriculture was for ancestors of contemporary societies. Using this measure, I estimate a positive relationship between the traditional use of irrigation and collectivism today, that is robust to instrumenting for the actual adoption of irrigation with the suitability of the local environment for irrigation. Chapter 3 then examines whether irrigation affected internal cultural norms in a sample of immigrants. I find that first and second generation immigrants originating from societies in which irrigation was used traditionally still have stronger collectivist beliefs. This result suggests that cultural transmission is an important mechanism for the long-run persistence of collectivism. Finally, I compare the degree of innovativeness of foreign-born migrants living

in the United States, to test whether pre-industrial agricultural techniques have long-term consequences for economic outcomes. Consistent with the idea that collectivism impedes innovation, I find that migrants whose ancestors practiced irrigated agriculture are less innovative.

CLIMATE RISK, COOPERATION, AND THE CO-EVOLUTION
OF CULTURE AND INSTITUTIONS[2]

At least since Charles-Louis De Montesquieu (1970 [1748]), social scientists have argued that climatic conditions affect human behavior and activity. In Chapter 4, co-written with Ruben Durante (Sciences Po), we examine the effect of a particular climatic dimension—the year-to-year variability of rainfall and temperature—on human cooperation. Climatic fluctuations posed a major source of risk for agriculture in pre-industrial societies. We hypothesize that the need of farmers to cope with climate risk and the experiences of mutual insurance favored the emergence of cultural norms of trust. As documented by ample historical evidence, many of the strategies to cope with weather-related risk involved cooperation of farmers with individuals from other communities, and oftentimes with strangers. A culture of greater trust could have developed in regions characterized by more variable weather which persisted over time, even after climate had become largely unimportant for economic activity.

We test this hypothesis in the context of Europe. To assess the degree of historical climate variability we use reconstructed paleoclimatic data for the period 1500–2000 that we combine with contemporary survey data at the sub-national level. We document a positive effect of year-to-year fluctuations in precipitation and temperature on trust today, conditional on a range of other geographic characteristics and country fixed effects. Consistent with our hypothesis, this result is mainly driven by variability during the growing season and by historical rather than recent variability. We then test for one mechanism that links climate and social cooperation. Using newly constructed data on the location of Medieval European land trade routes and major fairs, we show that regions with relatively higher climatic volatility were more closely connected to the network of Medieval trade. This result supports that inter-community exchange was an important coping strategy against climate risk. Finally, we investigate how cooperation persisted in the long run. Our empirical results suggest that cooperation enforcing local political institutions—measured by the presence of communal cities between 800 and 1800—emerged in regions with relatively stronger climatic fluctuations. Differences in local institutional quality persisted over time, since the same regions have better functioning sub-national governments even today. Taken together, Chapter 4 shows that pre-industrial climate risk has impacted social cooperation long-lastingly, by favoring the development of complementary, and mutually-reinforcing norms and institutions.

JOHANNES C. BUGGLE, *University of Lausanne*

---

[2] This chapter builds on and extends previous research by Durante (2009).

REFERENCES

Boyd, Robert, and Peter J. Richerson. *The Origin and Evolution of Cultures.* Oxford: Oxford University Press, 2004.

Buggle, Johannes C. "Law and Social Capital: Evidence from the Code Napoleon in Germany." *European Economic Review* 87 (2016): 148–75.

Durante, Ruben. "Risk, Cooperation and the Economic Origins of Social Trust: An Empirical Investigation." Mimeo, 2009.

Gerschenkron, Alexander. "Agrarian Policies and Industrialization: Russia 1861–1917." In *The Cambridge Economic History of Europe*. Cambridge: Cambridge University Press, 1966.

Gorodnichenko, Yuriy, and Gerard Roland. "Individualism, Innovation, and Long-Run Growth." *Proceedings of the National Academy of Sciences* 108, no. S4 (2011): 21316–19.

Montesquieu, Charles-Louis De. *The Spirit of the Laws*. New York: Simon and Schuster, 1970 [1748].

Tabellini, Guido. "The Scope of Cooperation: Values and Incentives." *Quarterly Journal of Economics* 123, no. 3 (2008): 905–50.

Triandis, Harry C., and Michele J. Gelfand. "A Theory of Individualism and Collectivism." *Handbook of Theories of Social Psychology* 2 (2012): 498–520.