Planter Elite Persistence and Long-Run Development in the South

Vafa Behnam Roudsari

- *"One reads the truer deeper facts of Reconstruction with a great despair. It is at once so simple and human, and yet so futile. There is no villain, no idiot, no saint. There are just men; men who crave ease and power, men who know want and hunger, men who have crawled. They all dream and strive with ecstasy of fear and strain of effort, balked of hope and hate. Yet the rich world is wide enough for all, wants all, needs all. So slight a gesture, a word, might set the strife in order, not with full content, but with growing dawn of fulfillment. Instead roars the crash of hell..."*

– W.E.B DuBois (1935)

Quick Summary

- The Census Bureau is announcing the release of a full 100% population count of the 1850, 1860 and 1870 Censuses in late September 2019. These are extremely valuable sources of information with a huge sample size.
- Want to see immediately what socio-economic mobility looked like in the immediate years surrounding the Civil War in the state of Alabama, considered a bedrock of Confederacy support, and to exploit heterogeneity of this measure.
- I then, more suggestively, want to examine whether this had any longer-run impacts on industrialization in the South.

Intellectual Forefathers

- Jonathan Weiner published an influential case-study in the 1970s examining the planter persistence rate in five adjacent Black Belt Alabama counties. He found there was a very high persistence rate, even after the war, approaching the persistence rates of Boston lawyers and doctors.
- This is in stark contrast to the findings of giants in Southern history, such as C. Vann Woodward (1951) or Randall & Donald (1969), who argued the once proud aristocratic class dissipated after the war.
- In many senses, my first-stage is a high-tech, “big-data” version of Weiner’s study, leveraging automated statistical linkage techniques and new fully digitized population censuses, a far larger sample than Weiner originally used.

Civil War

- The Civil War wrought incredible damage to the infrastructure of the South, whether financial or physical.
- Emancipation by itself liquidated an estimated $200,000,000 worth of assets (Garner 1964), far away the most valuable asset, and the asset that were the primary security used to backstop IOU’s in Southern finance (Woodman 1968).
- The Civil War and Reconstruction which ensued provided a huge opportunity to restructure the socio-economic relations in the South. (Foner 1988; Syrett 2005)

Reconstruction’s Biggest Failure

- Despite all the disagreements surrounding Southern history, the one consensus generated was that the most effective guarantee of the welfare of the newly freedmen would be a radical land transformation (Byres 1983; DuBois 1935; Ransom & Sutch 1977; Wright 1986). Not only would this have provided the freedman independence from the rigid labor markets of the South (Wright 1986), but it would have directly disabled the economic, and therefore political, base of their biggest adversaries, the planter class, who were the bedrock of the erstwhile Confederacy.
- With Andrew Johnson’s ascension to the Presidency in 1965, these hopes were all but dashed in the most idealistic of thinkers and Radical of Republicans. The military and political justifications for land reform that would have made it palatable to whites was not exploited, and America closed its doors to true black self-dependence.
- As a result, while certainly the freedmen were better off than they were under slavery, their conditions were awfully reminiscent of slavery, geographically immobile, and working as sharecroppers or laborers on land owned by others and tools leased by others and, in most cases, with the fruits of their labor appropriated by others (Ransom & Sutch 1997; Wright 1986).

Reconstruction and post-Reconstruction

- While there was no widespread land distribution, much debate still exists on whether the slaveholding planters in the South persisted in the post-War period, a time of massive social, political and economic change.
- Utilizing a full population count for the 1850, 1860 and 1870s released by the Census Bureau, and statistical record linkage techniques, I can try to create a longitudinal dataset to shed light on these processes.

Why Is This Investigation Important? Historical Perspective

- Still a big debate within Southern historiography: Did the slaveholding class in the South transition into a landholding elite post-war, or was there a "revolution in land titles"?
- If this planter class persisted, did they welcome industrialization, like the aristocratic elite did in "Old Prussia" and Japan (Billings 2011; Moore 1966), or were they a backwards-looking, pre-capitalist class (Ransom & Sutch 1977)?
- Furthermore, what did socio-economic mobility look like in the South in the immediate post-war era?
- Were the foundations for the "New South" set in the immediate post-Civil War era?

Why Is This Investigation Important? Economics Perspective

- What effects does this have for long-run development?
- Does having a long-standing planter elite make it easier for them to enact extractive policies and institutions and prevent the implementation of goods such as mass education? Is industrialization (where it occurs) limited in scope and made in the image of these planters?
- There is now a steady economics literature that stresses not only the relative power of elites, but also the particular **type** of elite.
- For example, Galor and Vollrath (2009) argue aristocratic, landholding elites stymie the development of human capital institutions. Rajan & Ramcharan (2011) argue that landholding elites restrict access to alternative sources of financing. Alston & Ferrie (1985) argue the development of the welfare state in the South was limited because planters wanted to “induce loyalty [in their workforce] through the provision of various in-kind benefits”.

Who Is a Planter Elite?

- No standardized definition, but usually meant to be minority of the white population who owned sizeable portions of land and slaves, and who were thought to have wielded immense political power, at county, state and federal level.
- We define persistence across two dimensions: wealth and geographic. I believe enforcing geographic persistence enforces a more accurate of the amount of control a wealthy planter family (as a corollary, it also makes the statistical record linkage easier, as we'll see in later slides)
- The only wealth variables available are value of real property and number of slaves.
- Peter Kolchin (1993) defines a planter as anyone who owns more than twenty slaves. Fogel & Engerman (1974) define large planters as those owning over 50 slaves, and medium planters as those owning between 16 and 50 slaves.
- However, we decided to go with real property value because the quality of slaves is greatly heterogeneous (Bonner 1944), and this variable obviously disappears in the 1870 census due to emancipation.
- We therefore use Weiner's real- property definition: all planters with at least $10,000 in real estate in 1850, $32,000 in 1860, and $10,000 in 1870.

Why Planter Persistence as Opposed to Standard Inequality Metrics?

- Why not just use income/wealth data and construct standard inequality indices, such as gini index, or 0.90-0.10 spread?
- I argue "planter persistence" provides three signals that standard inequality metrics do not: socio-economic mobility, class identity and relative power.
- I: These indices will not inform us how that class arose, i.e. through persistence and inheritance, or through "creative destruction". This has both economic and historical significance. In so far as social or economic mobility affects development and growth, this ought to have an effect on Southern economic outcomes.
- II: planter persistence can be a better proxy for relative political power. Ceteris paribus, we would expect a more persistent elite to have more ability and power in enacting vagrancy laws, limit rural-to-urban migration, limit the entry of competitors and limit public education schemes. They can also have an easier time enacting extractive institutions, with all its assorted negative outcomes for development (Acemoglu & Robinson 2012)

Why Planter Persistence as Opposed to Standard Inequality Metrics? II

- III: In Southern history there is a big debate on what or who constituted the post-civil war elite in the South, and whether this had an important impact on industrialization. Did the pre-war aristocratic planter class retain their power? And if they did, were the planter elite pro-industrialization (Billings 2011), or an anti-industrialization pre-capitalist class (Weiner 1975)? Furthermore, debate remains on whether this planter elite historians in the South offer anecdotal evidence there was a class conflict between merchants and the landholding class (Weiner 1975).
- Similarly, in political sociology and historical comparative development, Barrington Moore’s famous argument (1966) that the composition of elites has a very important effect on the path of industrialization taken. Dwight Billings adapted this to the Southern context, and argued that the South took the “Old Prussian” road to industrialization, one that was limited and circumscribed by the aristocratic elite.

Answering These Questions

- As intimated before, much debate still surrounds these issues.
- With the release of the 1860 and 1870 fully digitized population census, we can now statistically examine whether “elite” planters in 1870 were also “elite” in 1860. We can then compare this rate to the persistence rate from 1850-1860.
- Of course, while this is a crude and noisy measure, and could mask a lot of heterogeneity, as far as I know, this could very well be the ***first systematic*** investigation of the persistence rate in Alabama because of the new digitized population manuscripts.

Matching Records

- We link planters across the 1850-1860 and 1860-1870 Censuses.
- We need to standardize and clean names. Then, we feed these names, as well as other characteristics into an algorithm that predicts whether or not two individuals are the same individual. Using this we construct our persistence index.
- We use the procedure that Mills, Abramitzky and Perez (2019) described in their NBER working paper, that I'll spend the next few slides describing. One of the primary reasons for using their model is that it is unsupervised, i.e. does not require training data.

Why Is Matching Records Very Difficult?

- First order reasons: Under-enumeration, death and the low probability of being sampled in two independent, distinct surveys.
- Measurement errors: Mistakes in spelling names, birthplaces, ages, etc. or in the digitization of manuscript records.
- Therefore, matching individuals is not (or should not be) a straightforward, deterministic process.
- The approach we take here is a probabilistic one, leveraging a well-known statistics/machine learning technique called Expectation Maximization. We prefer a probabilistic approach because we would throw away too many data points if we enforce strict matches. If we could quantify the probability of being a match, we can fine-tune the Type I vs. Type II error tradeoff, and decide how conservative or liberal we want to be.

Why Use Expectation Maximization? – A Brief Foray into Statistical Record Linkage Theory

- The Census Bureau actually already has some linked datasets. Unfortunately, as of writing (7/21/2019), these only connect the Census year 1880 to all the other years. I could technically use the 1850-1880, 1860-1880 and 1870-1880 datasets, but this would clearly be an unrepresentative sample, which would lead to an upward bias in the persistence rate in years closer to 1880. Therefore I have to do the linking myself.
- Feigenbaum (2016) has a general, probabilistic method for linking datasets. This method however is supervised, i.e., it requires clearly and professionally labeled training data. Due to training data constraints, I must go for an unsupervised learning procedure.
- Mills, Abramitzky & Perez (2019) suggested a variant based on the statistical/machine learning tool Expectation Maximization. I follow their matching method, and describe it in the next few slides.

Setting up the Learning Problem

- Our fundamental data structure is a tuple (**a**, **b**), where **a** and **b** are two vectors consisting of matching characteristics, such as ages, names and place of birth.
- In general, we prefer to use "immutable" variables like age, name and place of birth, as opposed to "mutable" variables like occupation and place of residence, which can change within the lifespan of an individual.
- We compute scores for expected ages, names and place of birth.
- To measure the distance between two names, we use the Jaro-Winkler string distance measure. It has a lot of favorable intuitive properties, such as it punishes deviations in the beginning of names more harshly than those which occur towards the end of the names
- We then combine these into one score or probability.
- The Expectation Maximization algorithm computes a probability for each tuple of being in one of two clusters (matched or non-matched).
- In other words, the probability is a score of how close reported expected ages, names and place of birth are between two different individuals in the Census.

Expectation Maximization I

- We assume these distances are distributed by one of two clusters (i.e. multivariate distributions) one for the matches, and one for the non-matches.
- This particular implementation is parametric, so we assume some parametric distribution, usually Gaussian, for the two clusters. A Gaussian is fully determined by its mean and standard deviation.
- Therefore, our problem is just finding the parameters of the two multivariate Gaussians, mu_M, sigma_M, mu_U and mu_U, plus the unconditional probability of being a match, p_M (since there are only two events the unconditional probability of not being a match is p_U = 1-p_M)
- EM originally arose as a way to estimate models with latent variables, but we can consider the state of a given pair being a match or non-match as a latent variable as well

Expectation Maximization II

- We ultimately seek to learn $$\Pr(i \in M \,/\, \gamma_i) = \frac{\Pr(\gamma_i \,/\, i \in M)\, p_M}{\Pr(\gamma_i \,/\, i \in M)\, p_M + \Pr(\gamma_i / i \in U)\,(1 - p_M)}$$

(LHS: The probability of tuple i being a match (M), given gamma_i, the vector of distances)

- We parametrize Pr(y_i| i /in M) to a distribution. We usually assume a Gaussian, so the distribution has two parameters mu_M and sigma_M.
- If we had known which datapoints were matches and which weren't, this would be a maximum likelihood estimation problem and we could learn the four parameters mu_M, sigma_M, mu_U and sigma_U, as well as the unconditional probability p_M of being a match.
- However, we don't, so we begin with an initial guess for all five parameters.

Expectation Maximization III – E Step

- The EM algorithm basically consists of two steps: E and M step.
- After initializing all the parameters, we commence with the E step.
- E step: We compute the posterior probability at iteration t, defined as

$$\Pr\left(i \in M / \gamma_i, \theta_M^{(t)}, p_M^{(t)}\right)$$

$$= \frac{\Pr\left(\gamma_i / \theta_M^{(t)}\right) p_M^{(t)}}{\Pr\left(\gamma_i / \theta_M^{(t)}\right) p_M^{(t)} + \Pr\left(\gamma_i / \theta_U^{(t)}\right) (1 - p_M^{(t)})}$$

- Called E-step because this is the expected value of what we assumed to be the posterior distribution

Expectation Maximization IV – M Step

- M-step: Using the posterior computed in the E step, we now compute updates for all five parameters (using theta as a shorthand for the two parameters, i.e. theta = (sigma, mu))

$$p_M^{(t+1)} = \frac{1}{N}\sum_{i=1}^{N} w_i^{(t)}$$

$$\theta_U^{(t+1)} = arg\max_{\theta}\sum_{i=1}^{N}(1 - w_i^{(t)})\, log\, \Pr\,(\gamma_i/\theta)$$

$$\theta_M^{(t+1)} = arg\max_{\theta}\sum_{i=1}^{N} w_i^{(t)}\, log\, \Pr\,(\gamma_i/\theta)$$

- Repeat the E and M-step until some stopping criterion.
- Each iteration is guaranteed to improve the empirical likelihood score, but it is not guaranteed that this will be the global maximum, nor will the rate at which it converges to this critical point be reasonably quick.

Expectation Maximization V – Decision Rules

- We now have probabilities for two individuals being matches. Now, we need to make some (ad-hoc) thresholds to judge what is and what isn't a match.
- "Formally, this decision rule can be stated in the following way: To be considered a unique match for a record in dataset A, a record in dataset B has to satisfy three conditions. Specifically, the researcher should: 1. choose the match with highest probability of being a true match out of all potential matches for the record in A. 2. choose a match that is true with a sufficiently high probability, i.e. a match with a probability p1 that satisfies p1>p for a given p in (0; 1] chosen by the researcher. 3. choose a match for which the second best match is unlikely, i.e. the match score of the next best match, denoted as p2, satisfies p2"
- Adjusting these decision rules is akin to manipulating the tradeoff between the Type 1 and Type 2 errors. Because our sample is small, I use more liberal decision rules to increase the sample as much as possible. This is more acceptable in this context because we only care for matches within the same county (which makes a false positive less likely)

Statistical Specification

$$Y_{ct} = B_t \Delta Pers\_rate_c + \lambda_t + \lambda_c + \theta_t X_c + \epsilon_{ct}$$

Y_{ct} is an outcome at county c at time t. $\Delta Pers\_rate_c$ is the difference in the persistence rate in 1860-1870 and 1850-1860. λ_t and λ_c are timed and county fixed effects respectively, while $\theta_t X_c$ are time-varying county controls. We will discuss standard errors in the next slide.

We also do a 2SLS, using what I argue is an exogenous variation from Wilson's Raid (to be explained in later slides). The second stage is as above, with the first stage of this equation being:

$$\Delta Pers\_rate_c = D\mathbb{1}\{Wilson'sRaid\}_c + \epsilon_c$$

Where Wilson's Raid is a dummy equal to 1 if the county was invaded by Wilson's riders in 1865, and 0 if not.

Standard Errors

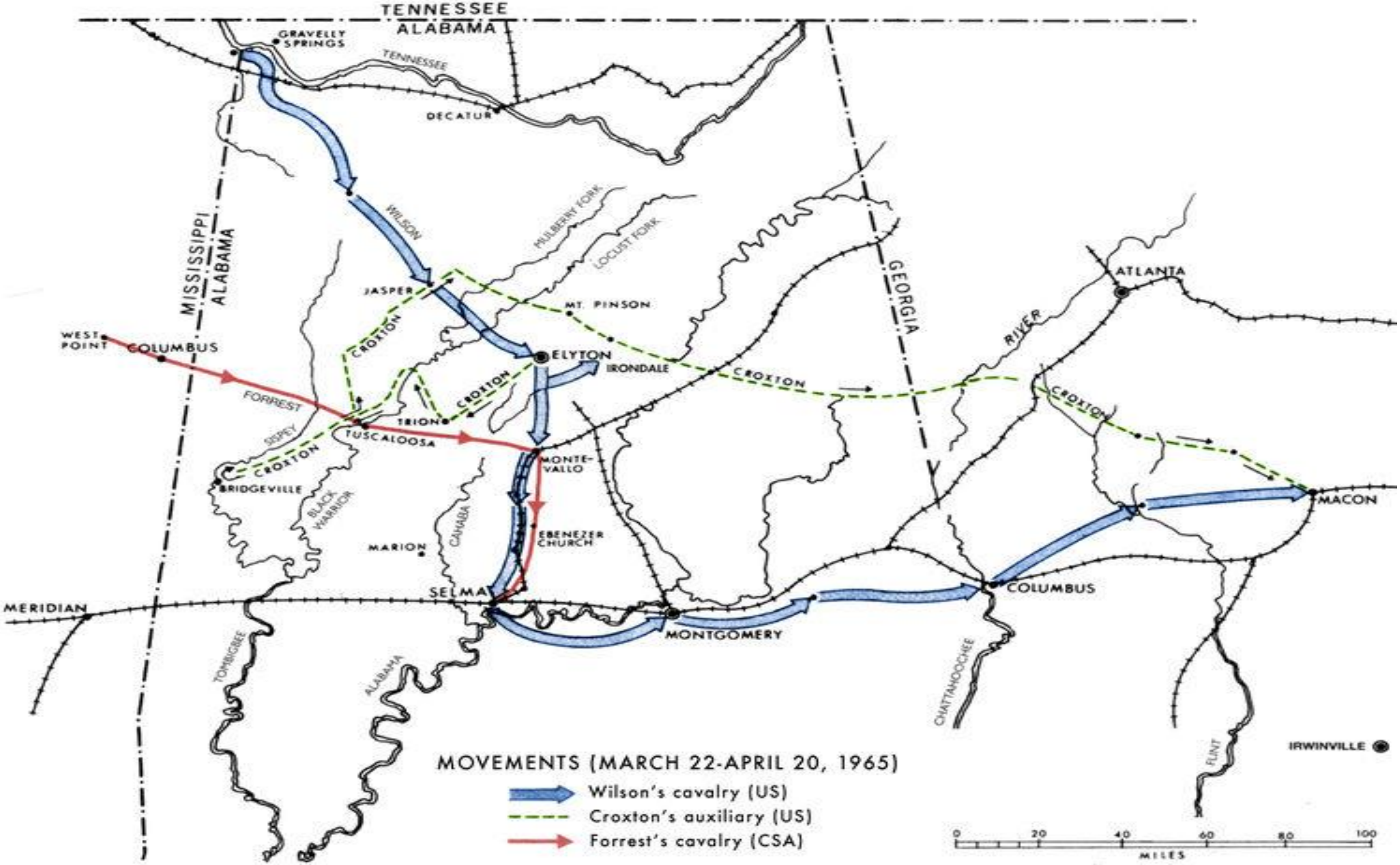
- I use Huber-White robust standard errors (which clusters at each state x time).
- But to account for the possibility of serial autocorrelation, I also cluster at the county level alone, which allows correlation within counties across time. We have enough clusters to do cluster only at the county level safely (50 counties is enough).
- I also cluster at the congressional district level, a higher level of aggregation than counties, to account for spatial correlation between neighboring counties. Since, by advice of Angrist & Pischke (2008), the number of clusters in this case is “small”, we utilize a t-distribution for hypothesis testing on (#congressional district – 1) degrees of freedom (Land and Donald 2007)

Dependent Variables

- I utilize a number of industrial and agricultural dependent variables. These are designed to capture numerous dimensions of the rate of industrialization or labor-saving technologies, investment in human capital and socioeconomic distribution.
- % of (sharecroppers + tenants)/(wage laborers + independent land owners) is a proxy for paternalistic relationships in the South, as well as geographic mobility (Alston & Ferrie 1985)
- Year in which (machine harvested cotton)/(total harvest) exceeded 10%
- However, noting that the diffusion mechanical cotton picker happened primarily in the WW2 period, we can also use, as Naidu and Hornbeck (2012) did, value of agricultural capital equipment, machinery as well as labor-saving “technology” horses and mules as a proxy for “mechanization”
- Enrollment in school and literacy rates in school

Seeking out a Natural Experiment

- As well as the fixed effects design of the study, I have access to the 1850, 1860 and 1870 agricultural schedules, also conducted by the Census Bureau, which contains a rather impressively broad range of agricultural controls, which I include
- Of course, some may still be concerned that there is omitted variables in this study. One way to solve for this is to leverage a natural experiment.
- Naidu and Hornbeck (2014) use flooding and black out-migration in Mississippi to analyze rates of mechanization.
- I argue Wilson's Raid at the tail end of the war meets the criteria of a natural experiment.


TENNESSEE
ALABAMA
GRAVELLY SPRINGS
TENNESSEE
DECATUR
WILSON
MULBERRY FORK
LOCUST FORK
MISSISSIPPI
ALABAMA
GEORGIA
JASPER
MT. PINSON
ATLANTA
RIVER
WEST POINT
COLUMBUS
CROXTON
ELYTON
IRONDALE
FORREST
SIPSEY
TRION
TUSCALOOSA
MONTE-VALLO
BRIDGEVILLE
BLACK WARRIOR
CAHABA
MARION
EBENEZER CHURCH
MACON
SELMA
COLUMBUS
MERIDIAN
MONTGOMERY
TOMBIGBEE
ALABAMA
CHATTAHOOCHEE
FLINT
IRWINVILLE
MOVEMENTS (MARCH 22-APRIL 20, 1965)
Wilson's cavalry (US)
Croxton's auxiliary (US)
Forrest's cavalry (CSA)
0 20 40 60 80 100
MILES

Wilson’s March

- Near the end of the war, two weeks before Robert Lee surrendered his army, Commander Wilson marched through Alabama to destroy the last remaining munitions factories and to dispense of the last few Confederate forces.
- Wilson was ordered to destroy the last remaining war plants in Alabama. From his starting position in the Tennessee River, Wilson took basically the quickest path southwards to the industrial centers in Selma and Montgomery (McMurry 1977).
- The Jonathan Weiner study is suggestive that regions on the path of Wilson’s Raid had lower rates of planter persistence. However, since Weiner did his study in the 70s without such high-level computing power, we can do his same analysis but with much more power.

Data Improvements

- The first and biggest is the release of the full dataset. The Census Bureau is planning to release a fully digitized version of the 1860 and 1870 Population Census.
- As Steven Ruggles, founder of the iPUMS database states, “The Church of Jesus Christ of Latter-Day Saints (LDS) and the Minnesota Population Center (MPC) have produced a remarkable machine-readable database encompassing the entire population of the United States enumerated in the 1850, 1860 and 1870 Census”
- In the meantime, I tried to do this analysis with the already published Census data. Currently we only have the 1860 and 1870 1-in-100 sample available, as opposed to a full 100% population count.
- I tried to do this analysis with these already available datasets, and with different decision rules, but the study was too underpowered to even be suggestive. (Makes sense, because there is a 1-10000 chance that an individual will be in both surveys, even assuming no death, under-enumeration or transcription errors).

Main Goals of the Study

- Show that elite persistence reveals more about the political economy of a region than standard inequality metrics.
- Leverage a new and massive dataset, and use/adapt a statistical algorithm, Expectation Maximization, in the particular context of matching census records, even without training data.
- Develop a quick understanding of the structure and main players of the Southern economy in the immediate post-war period.
- Understand what effects such a structure had on long-run economic outcomes such as industrialization, financialization and mass education by exploiting heterogeneity in this structure.

References I

- Donald, Stephen G. and Kevin Lang (2007) *Inference with Difference-In-Differences and Other Panel Data*. Review of Economics and Statistics
- Sutch, R., & Ransom, R. (1977). *One Kind of Freedom.* Cambridge University Press.
- DuBois, W. (1935). *Black Reconstruction in America, 1860-1880.*
- C. Vann Woodward (1955), *Origins of the New South, 1877-1913*
- J. G. Randall and David Donald (1969), *The Civil War and Reconstruction.*
- Weiner, J. (1976). Planter Persistence and Social Change: Alabama, 1850-1870. *The Journal of Interdisciplinary History*.
- Garner, J. (1964). *Reconstruction in Mississippi.*
- Foner, E. (1988). *Reconstruction: America's Unfinished Revolution, 1863-1877.* Harper Perennial Modern Classics.
- Syrett, J. (2005) *The Civil War Confiscation Acts failing to reconstruct the South*
- Wright, G. (1986). *Old South, New South: Revolutions in the Southern Economy Since the Civil War.*
- Byres, T. (1983). *Sharecropping and Sharecroppers.*
- Billings, D. (2011). *Planters and the Making of a "New South“, Class, Politics, and Development in North Carolina, 1865-1900*
- Moore, B. (1966) *Social Origins of Dictatorship and Democracy*

References II

- Galor, O. & Vollrath, D. (2005). *Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence.* Brown University Economics Working Paper No. 06-14
- Rajan R. & Ramcharan R. (2011). *Land and Credit: A Study of the Political Economy of Banking in the United States in the Early 20th Century.* The Journal of Finance.
- Alston, L. & Ferrie, J. (1985) *Labor Costs, Paternalism, and Loyalty in Southern Agriculture: A Constraint on the Growth of the Welfare State.* The Journal of Economic History Vol. 45, No. 1
- Kolchin, P. (1993). *American Slavery, 1619-1877.*
- Fogel, R. & Engerman, S. (1974). *Time on the Cross: The Economics of American Negro Slavery.*
- Bonner, J. (1944). *Profile of a Late Ante-Bellum Community*, American Historical Review.
- Acemoglu, D. & Robinson, J. (2012). *Why Nations Fail.* Crown Publishing Group.
- Feigenbaum, J. (2016). *A Machine Learning Approach to Census Record Linking.* Working Paper.
- Abramitzky R., Mills R. & Perez S. (2019) Linking Individuals Across Historical Sources: A Fully Automated Approach. NBER Working Papers.
- Angrist, J. & Pischke, J. (2009). *Most Harmless Econometrics: An Empiricist's Companion.*
- Naidu, S. & Hornbeck, R. (2014). *When the Levee Breaks: Black Migration and Economic Development in the American South.* American Economic Review, Vol 104, No.3.
- Pickett Jones, J. (1977). *Yankee Blitzkrieg: Wilson's Raid Through Alabama and Georgia.*
- Ruggles, S (2002). LINKING HISTORICAL CENSUSES: A NEW APPROACH. History and Computing.