Appendix A: Definition of Free/Slave Boundary and Spatial Extent of Data

We use standard definitions of the U.S. states where slavery was legal in 1860. This excludes territories, e.g. Kansas and Nebraska. We classify as 'free' those states, e.g. New Jersey and Illinois, where general emancipation had taken place well before 1860, but there remained some former slaves bonded under transitional indentureship, for example. This gives the following 'slave' states: Alabama, Arkansas, Delaware, the District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, North Carolina, South Carolina, Tennessee, Texas, and Virginia. This set of states gives a clearly defined border that separates the country into two sections, one slave and one free. The resulting free/slave boundary is defined, from west to east, as follows:

- The Missouri/Iowa border
- The Missouri/Illinois border, which largely follows the Mississippi River down to Cairo, Illinois
- The northern border of Kentucky from Cairo, Illinois, to Ashland, Kentucky, which largely follows the Ohio River
- The northern border of (West) Virginia, along the Ohio River
- The western border of Pennsylvania with the northern (West) Virginia panhandle
- The southern border of Pennsylvania with (West) Virginia and Maryland, which follows the Mason-Dixon Line
- The Delaware/Pennsylvania border
- The midline of the Delaware River between New Jersey and contiguous Delaware.

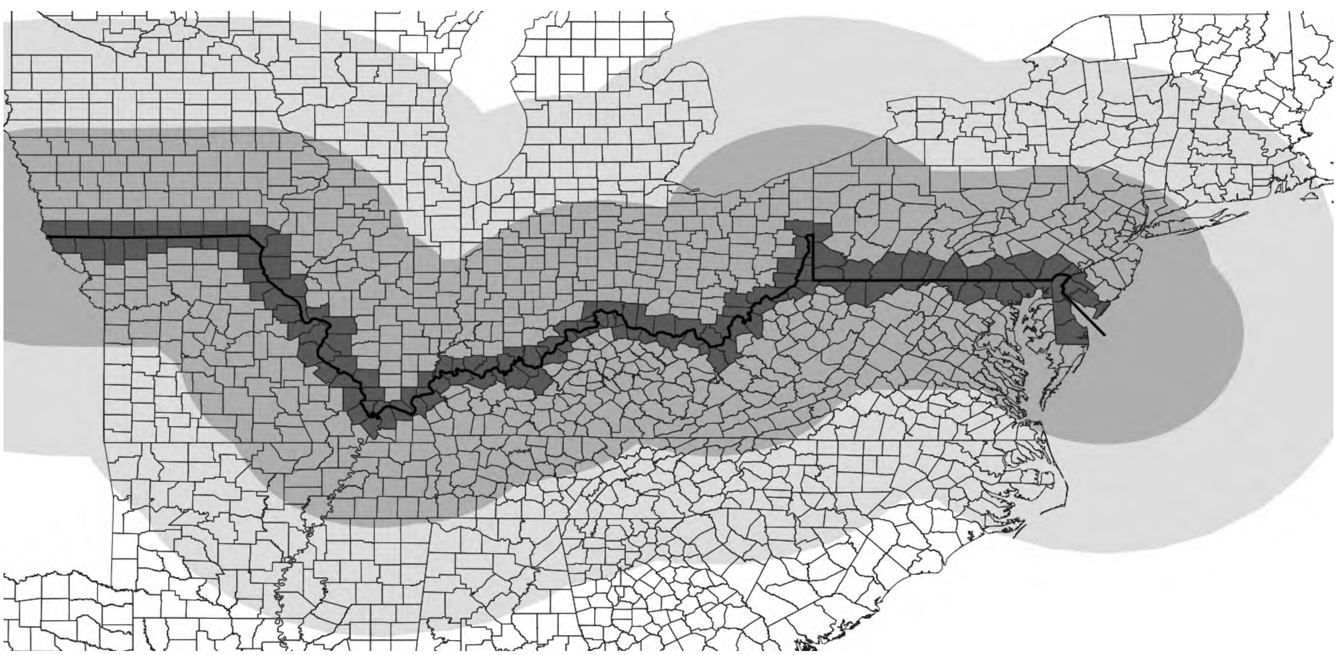
We use spatial data from the NHGIS project (Minnesota Population Center, 2011) to map the free/slave boundary and to measure counties' proximity to said boundary. We present this boundary in Figure 1 of the paper and here in Appendix Figure A.1. To this map, we add the 1860 county boundaries, per NHGIS, for reference. In the paper, we use two distinct concepts of proximity: adjacency and distance. Adjacency refers to a county directly touching the free/slave boundary. For example, the 1860 counties that are adjacent to the free/slave boundary are shown with dark-gray shading. (The proximity measures are computed separately for each year of data.) We also construct a buffer of 150 and 300 miles from the boundary. Counties adjacent to the boundary fall within these two buffers. Additional areas within the 150-mile buffer are shaded in medium gray in Appendix Figure A.1. In addition to those two, areas within 300 miles are shaded in light gray. Counties with any portion lying within these buffers are categorized in the relevant buffer zones. As a control variable, we also compute the distance from the border for each county's centroid and the *average* distance of a county to the free/slave border by computing distance to the boundary for a high-dimensional (10kx10k) raster over the contiguous US and then by calculating the average value within each county. As another control, we use the latitude and longitude of each county's centroid, as supplied by NHGIS.

The presence of riverine boundaries necessitates further discussion. Boundaries on rivers are typically defined on a specific side of a river, or perhaps at a midpoint. Changes in the course of a river over time or poor surveying at the time of setting the border might generate discrepancies between a boundary and the current course of a river, even to the point of generating exclaves. For example, Kaskaskia is an exclave of Illinois created by a change in the course of the Mississippi River. We rely not on the contemporary river course, but rather the NHGIS definition of the historical (1860) boundary to set state borders. For the most part, the rivers are sufficiently narrow and their historical meanders sufficiently small so as to make little difference for our classification of counties with respect to the free/slave boundary. Counties that are adjacent to an above-named river segment will be adjacent to the free/slave border as well, for example. The major exception is the Delaware River, which turns into a bay between Delaware and New Jersey. The statutory boundary between New Jersey and Delaware lies on the New Jersey side of the Delaware River, except for a small exclave of Delaware on the New Jersey side. We choose the midpoint of the river instead to better calibrate the measure of distance to the boundary. This does not change the adjacency concept for counties on either side of the Delaware River, but it brings the distance measurement into better balance between sides of the river.

References

Minnesota Population Center. *National Historical Geographic Information System: Version 2.0.* Minneapolis, MN: University of Minnesota 2011.

Appendix Figure A.1: The 1860 Free/Slave Boundary and Several Measures of Proximity



Notes: this map displays 1860 county boundaries (thin black lines), the free/slave boundary (thick black line), and three measures of proximity to said boundary. The counties that touch the free/slave boundary are shaded in dark gray. Additional areas that lie within a buffer of 150 miles from the boundary have medium gray shade. Further areas within a buffer of 300 miles from the free/slave boundary are denoted with a light gray shade.

Appendix B: Land Values and the Scarcity of a Variable Factor

The Basic Problem and a First-Order Solution: How does the price of the fixed input (e.g., land) respond to a change in price of a variable input (e.g., a type of labor)? To fix ideas, let inputs $x_i, i \in \{2...k\}$, be supplied perfectly elastically at fixed prices p_i. In contrast, input $i = 1$ is in fixed supply locally and has endogenous price p_1. Inputs are defined to be positive: $x_i \geq 0 \, \forall \, i$.

Assume that factor markets are perfectly competitive, including the land market. This implies zero profits in equilibrium. One could equivalently adopt a 'Ricardian' approach: land owners earn the residual (rent) once other factors have been paid. If there is an open market for land, then land rents at a location should be capitalized into the price of land. Therefore, profits earned on variable factors should equal the opportunity cost of land.

In response to a parameter change, the variable factors adjust, as does the price of the fixed factor. Let $\mathrm{d}p_2 > 0$ be given. What is $\mathrm{d}p_1$? Consider the profit function for a particular location: $\pi(\vec{p})$. This gives the maximum profits attainable for a given price vector. Taking full differentials of $\pi = 0$ yields

$$\pi_1 \mathrm{d}p_1 + \pi_2 \mathrm{d}p_2 = 0, \tag{*}$$

which gives the following equilibrium relationship:

$$\frac{\mathrm{d}p_1}{\mathrm{d}p_2} = -\frac{\pi_2}{\pi_1}$$

Recall Hotelling's Lemma: the derivative of the profit function w.r.t. p_i is factor demand $-x_i$. Therefore $\frac{\mathrm{d}p_1}{\mathrm{d}p_2} = -\frac{x_2}{x_1}$, which is negative (strictly so if $x_2 > 0$). In terms of elasticities,

$$\epsilon_{12} \equiv \frac{p_2}{p_1}\frac{\mathrm{d}p_1}{\mathrm{d}p_2} = -\frac{p_2 x_2}{p_1 x_1},$$

the negative ratio of the expenditure shares.

It is intuitive that the prices are related in this way. When p_2 changes, it affects the net worth of the firm. The firm/farm's net worth shrinks more if the expenditure share on x_2 is large. This decline in net worth is spread more widely if the expenditure share on x_1 is itself large.

Discussion of Returns to Scale: We did not state an assumption above about returns to scale. Nevertheless, this is implicit in invoking perfect competition. An assumption of a finite solution is inconsistent with simultaneously being a price taker and facing increasing returns to scale. So perhaps we should assume non-increasing returns. Further, a typical counterargument to a claim of decreasing returns to scale (DRS) is that instead there is some unspecified input (e.g., proprietor's labor). If we could properly catalog and duplicate all inputs, then this cloned set of factors would generate just as much output as the original. This analysis suggests that the leading case is constant returns to scale (CRS). Under constant returns, Euler's Theorem tells us that, if prices equal marginal revenue products, then the sum of factor payments equals total output. This implies that the land rent that we calculate as a residual also equals the price that would prevail in a competitive market for land.

If the production function does exhibit DRS, then Euler's theorem tells us that the residual rent would be different from the marginal product of land. What would equilibrium in the land market look like in this case? DRS also has the unrealistic implication that production units should be infinitesimally small. It is likely, however, that some indivisibility (e.g., farm operators coming in integer units) would prohibit farm units from being below some minimum scale. If so, the farm size would be constrained and therefore the optimal choice of land would not be determined by the first order condition. The market value of farms at this minimum scale would instead be pinned down by the profit (the residual implied by Euler's Theorem).

The example of Cobb-Douglas with constant returns: Assumptions are as above, except that output is determined by a Cobb-Douglas with constant returns to scale (CRS). The ith factor has share $\alpha_i \in (0,1)$. The FOCs define the relative factor inputs, e.g.,

$$\frac{p_1 x_1}{\alpha_1} = \frac{p_i x_i}{\alpha_i} \;\; \forall\, i = 1, ..., k.$$

As the first factor is fixed in size, we can define the others in terms of x_1:

$$x_i = \frac{p_1}{p_i}\frac{\alpha_i}{\alpha_1} x_1$$

The price of output is the numeraire. Output (Y) and cost (C) are as follows:

$$Y \equiv \prod_{i=1}^{k} x_i^{\alpha_i} = \prod_{i=1}^{k} \left(\frac{p_1}{p_i}\frac{\alpha_i}{\alpha_1} x_1\right)^{\alpha_i} = \frac{p_1 x_1}{\alpha_1} \prod_{i=1}^{k} \left(\frac{\alpha_i}{p_i}\right)^{\alpha_i}$$

$$C \equiv \sum_{i=1}^{k} p_i x_i = \sum_{i=1}^{k} p_i \left(\frac{p_1}{p_i}\frac{\alpha_i}{\alpha_1} x_1\right) = \frac{p_1 x_1}{\alpha_1} \sum_{i=1}^{k} \alpha_i$$

Per the CRS assumption, $\sum_{i=1}^{k} \alpha_i = 1$, so $C = p_1 x_1/\alpha_1$. Profit is as follows:

$$\pi = Y - C = \frac{p_1 x_1}{\alpha_1} \left(\prod_{i=1}^{k} \left(\frac{\alpha_i}{p_i}\right)^{\alpha_i} - 1\right)$$

Competitive markets for factors imply a zero-profit condition: $\pi = 0$, or $\prod_{i=1}^{k} (\alpha_i/p_i)^{\alpha_i} = 1$, or

$$p_1 = \alpha_1 \prod_{i=2}^{k} \left(\frac{\alpha_i}{p_i}\right)^{\frac{\alpha_i}{\alpha_1}}$$

This yields an elasticity equal to the (negative of the) ratio of the factor shares, as in the generic version above. (Take logs of both sides to see this.) For Cobb-Douglas, however, this is not an approximation, but rather an elasticity that holds globally, in logs, at any interior solution.[1]

We could also derive this price and elasticity from optimal land use (i.e., the demand curve for land as a productive factor). The FOC is

$$\begin{aligned} p_1 &= \alpha_1 Y / x_1 \\ p_1 &= \frac{\alpha_1}{x_1}\left(\frac{p_1 x_1}{\alpha_1}\prod_{i=1}^{k}\left(\frac{\alpha_i}{p_i}\right)^{\alpha_i}\right) \\ p_1 &= \alpha_1 \prod_{i=2}^{k}\left(\frac{\alpha_i}{p_i}\right)^{\frac{\alpha_i}{\alpha_1}} \end{aligned}$$

which is the same as the price derived from the zero-profit condition. This is as predicted by Euler's Theorem.

Second-order effects: Why does the general elasticity not depend on the extent of substitution between factors? Hotelling's Lemma is a first-order result that follows an application of the Envelope Theorem. A small change in price will occasion a set of small quantity changes. But the effect on profits of small quantity changes is approximately zero near the optimum, which is where the profit function is evaluated.

Accordingly, the dependence on factor substitutability only appears when considering second-order changes. The first derivative, $\mathrm{d}p_1/\mathrm{d}p_2$ is negative, so that the increase in p_2 decreases p_1. Does the second derivative amplify this effect, $\frac{\mathrm{d}^2 p_1}{\mathrm{d}p_2^2} < 0$, or attenuate it?

The zero-profit condition ($\pi(\vec{p}) = 0$) still holds, even for large changes, because p_1 adjusts to absorb any surplus generated by the other factors. Taking full differentials of the zero-profit conditions to first and second order yields the following equations:

$$[\nabla\pi(\vec{p})]\,\vec{\mathrm{d}p} = 0\;;\;\vec{\mathrm{d}p}'\left[\nabla^2\pi(\vec{p})\right]\vec{\mathrm{d}p} = 0$$

where $\vec{\mathrm{d}p} = [\ \mathrm{d}p_1\ \mathrm{d}p_2\ 0\ \ldots\ 0\ ]'$ and ∇ is the operator that takes derivatives w.r.t. the price vector $\vec{p}$. The first-order equation reproduces equation $(*)$ from above. The second-order equation has a matrix pre- and post-multiplied by the infinitesimal-change-in-price vector:

$$[\ \mathrm{d}p_1\ \mathrm{d}p_2\ 0\ \cdots\ 0\ ]\begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \cdots & \pi_{1k} \\ \pi_{21} & \pi_{22} & \pi_{23} & \cdots & \pi_{2k} \\ \pi_{31} & \pi_{32} & \pi_{33} & \cdots & \pi_{3k} \\ \pi_{41} & \pi_{42} & \pi_{43} & \cdots & \pi_{4k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \pi_{k1} & \pi_{k2} & \pi_{k3} & \cdots & \pi_{kk} \end{bmatrix}\begin{bmatrix} \mathrm{d}p_1 \\ \mathrm{d}p_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

[1] Under CRS, the zero-profit condition can be re-written as $\vec{\alpha} \cdot [\log(\vec{p}) - \log(\vec{\alpha})] = 0$. One set of prices consistent with an interior solution is $\vec{p} = \vec{\alpha}$. To stay in the zero-profit set, a deviation of some price p_i from this vector has to be matched by offsetting proportional deviations in other prices, with the elasticities defined by the relative factor shares.

Turn the crank once:

$$[\,\mathrm{d}p_1\ \mathrm{d}p_2\ 0\ \cdots\ 0\,]\begin{bmatrix}\mathrm{d}p_1\sum_{i=1}^{k}\pi_{i1}\\ \mathrm{d}p_2\sum_{i=1}^{k}\pi_{i2}\\ 0\\ \vdots\\ 0\end{bmatrix}=0$$

Turn the crank again:

$$(\mathrm{d}p_1)^2\sum_{i=1}^{k}\pi_{i1}+(\mathrm{d}p_2)^2\sum_{i=1}^{k}\pi_{i2}=0$$

which gives us

$$\frac{\mathrm{d}^2p_1}{\mathrm{d}p_2^2}=-\left(\sum_{i=1}^{k}\pi_{i2}\right)\bigg/\left(\sum_{i=1}^{k}\pi_{i1}\right) \tag{$\dagger$}$$

The second derivative is negative if the two sums have the same sign. Do they? Recall that the profit function has a Hessian ($\mathbf{H}\equiv\nabla^2\pi$) that is negative semi-definite.[2] Therefore, for any non-zero vector w,

$$w'\mathbf{H}w\leq 0. \tag{$\ddagger$}$$

Consider the two vectors $w_1=[\,1\ 0\ 0\ \cdots\ 0\,]'$ and $w_2=[\,0\ 1\ 0\ \cdots\ 0\,]'$. Notice that using w_i in the inequality ($\ddagger$) constructs the ith column sum. This allows us to rewrite equation ($\dagger$) as follows:

$$\frac{\mathrm{d}^2p_1}{\mathrm{d}p_2^2}=-\frac{w_2'\mathbf{H}w_2}{w_1'\mathbf{H}w_1},$$

which, by ($\ddagger$), is weakly negative (strictly so, as long as $w_2'\mathbf{H}w_2$ is not zero).

Because the second derivative (of p_1 w.r.t. p_2) is negative, allowing for substitution among the x_i when $p_2\uparrow$ amplifies the reduction in p_1, as compared to the first-order effect. Intuitively, there are contrasting parts of the second-order effect of $p_2\uparrow$. If the first two inputs are substitutes, the decrease in x_2 shifts out the demand for x_1, *ceteris paribus*. But the decline in profits also spurs the flight of other mobile factors from the area. This latter effect dominates the former one for any interior solution at which we can take derivatives.

The case of perfect substitutes: Here substitutability is cranked up to the max, and the relevant functions are not differentiable at an interior solution. If all factors are perfect substitutes, then we can write output as $Y=\sum_{i=1}^{k}\alpha_i x_i$. This case is, in fact, quite uncomplicated. The price of land is independent of the other factors. The marginal product of x_1 is α_1, which determines the price p_1. The effect of other factor prices on p_1 is nil. We have to careful with this case, though, because it is easy to scale up in spite of a fixed supply of land. If any of the other factors, $i>1$, have $p_i<\alpha_i$, then one could increase those factors and increase profits. The fixed factor is in no way a constraint to unlimited growth.

[2]Inputs were defined positively above. If we had defined a net-output vector instead, the Hessian would be positive semi-definite. What matters is the ratio (left-hand side of ($\dagger$), in which this sign convention cancels out.

Appendix C: Measurement of Glacial Extent

We digitized the location of the terminal moraine using maps published by George Frederick Wright (1884, 1890, 1892). This geological feature is a "well defined southern limit to the marks of glacial action in the United States" (Wright, 1884, page 203). We worked with the most detailed maps provided for each mapping segment. We also used the NHGIS files for 1860-1890 to help georeference points on Wright's maps.

Going from east to west, segments within a given area were digitized using the indicated maps.

- Massachusetts and New York: Wright, 1884, Plate 1 and text on page 203.
- New Jersey: Wright, 1884, plate 2.
- Pennsylvania: Wright, 1884, plate 3.
- Ohio: Wright, 1884, Plates 8 through 16.
- Kentucky: Wright, 1884, Plates 5, 16 and 17. The latter two plates were more detailed, but only covered the Cincinnati area. Plate 5 was used for the area around Madison, Indiana.
- Indiana: Wright, 1890, Figure 3.
- Illinois: Wright, 1890, Plate 5.
- Driftless region: Wright, 1892, unnumbered map, facing page 68.
- Missouri: Wright, 1892, text on page 96 and unnumbered map facing page 68.

Appendix Figure C.1 displays the terminal moraine (glacial boundary) as a dashed line. For comparison, the free/slave boundary is shown as a thick, solid line, and contemporary state boundaries are displayed as thin, black lines. In general terms, the glacial and free/slave boundaries are both oriented in an east/west direction, but they do not precisely coincide. The southern extent of the glacier is to the north of the free/slave boundary, with three exceptions. The greatest exception is in the state of Missouri, which is approximately split in half by the terminal moraine. This moraine crosses the Mississippi river near St. Louis, and generally follows the course of the Missouri River and then the Osage River. Somewhat downriver of St. Louis, the moraine is close to the Mississippi River, but stays on the Illinois side. (Of the areas with greatest slaveholding in Missouri, the 'Little Dixie' region is largely in the glaciated region, while the 'Bootheel' region is not at all.) The terminal moraine also cuts into Kentucky for a short stretch across the river from Madison, Indiana, and for a longer stretch across the river from Cincinnati, Ohio. (See Appendix Figure C.2, Panel A, for detail.) Away from these areas, the southernmost glacial extent cuts a path significantly to the north of the free/slave border. Apart from those noted above, the closest approach of the terminal moraine to the free/slave boundary is at the Wabash River and at the northern panhandle of West Virginia. (See Appendix Figure C.2, Panel B, for example.)

In some areas, the terminal moraine is superficially noteworthy as a ridge. In others, it is less noticeable to a casual, surface observer. In all areas, however, the extent of the glacier can be determined by the presence or absence of rock striations and glacial till, and other features well understood by geologists.

We also digitized the location of the 'Driftless Region,' an area north of the terminal moraine that was nevertheless not subject to glaciation. It is mostly found in Southwest Wisconsin. (See Appendix Figure C.2, Panel C, and note the rotation of the map, such that north points right.)

References

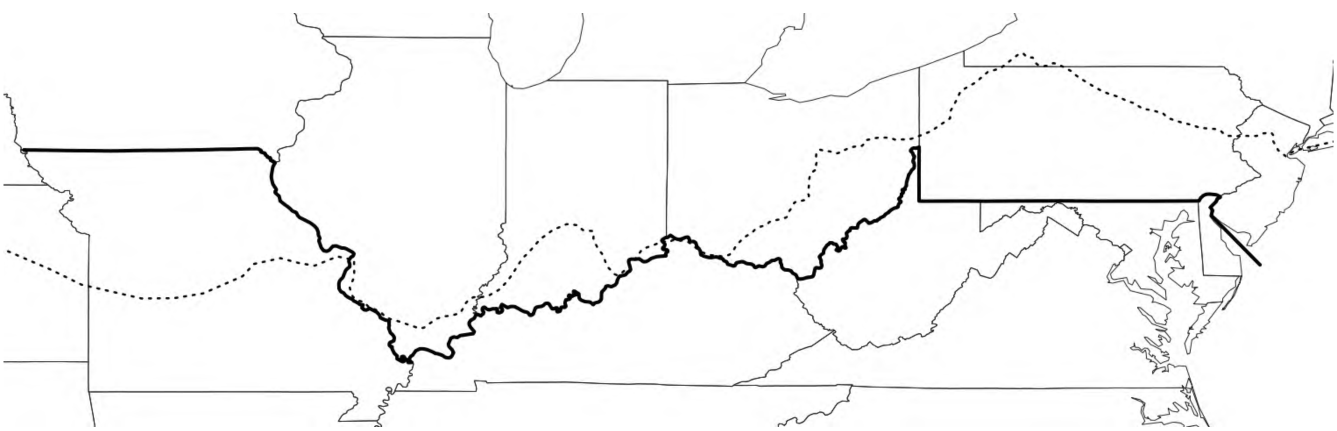
Minnesota Population Center. *National Historical Geographic Information System: Version 2.0.* Minneapolis, MN: University of Minnesota 2011.

Wright, G. Frederick (1884). *The glacial boundary in Ohio, Indiana and Kentucky.* Cleveland, Ohio: Western Reserve Historical Society.

Wright, G. Frederick (1890). *The glacial boundary in western Pennsylvania, Ohio, Kentucky, Indiana, and Illinois.* Bulletin of the United States Geological Survey ; no. 58. Washington, D.C.: Department of the Interior, United States Geological Survey.

Wright, G. Frederick (1892). *Man and the glacial period*. Akron, Ohio: The Werner Company.

Appendix Figure C.1: Terminal Moraine, as compared to Free/Slave Boundary and State Borders

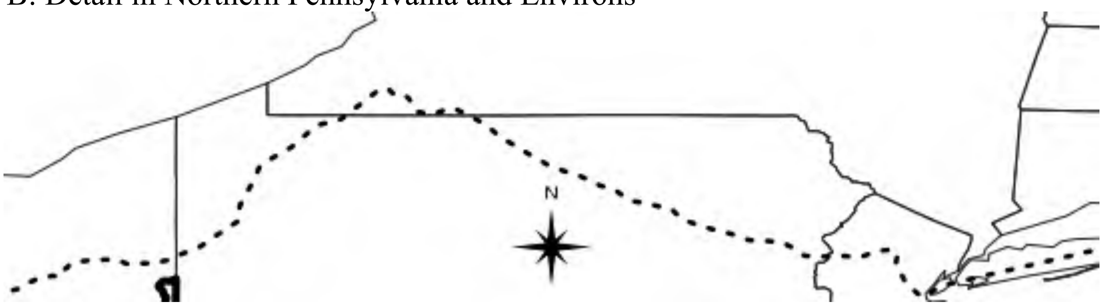


Sources: NHGIS (2011) for state boundaries, plus authors' calculations for free/slave boundary; Wright (1884, 1890, 1896) for terminal moraine, plus authors' digitization. See text of Appendix C for detailed sources.

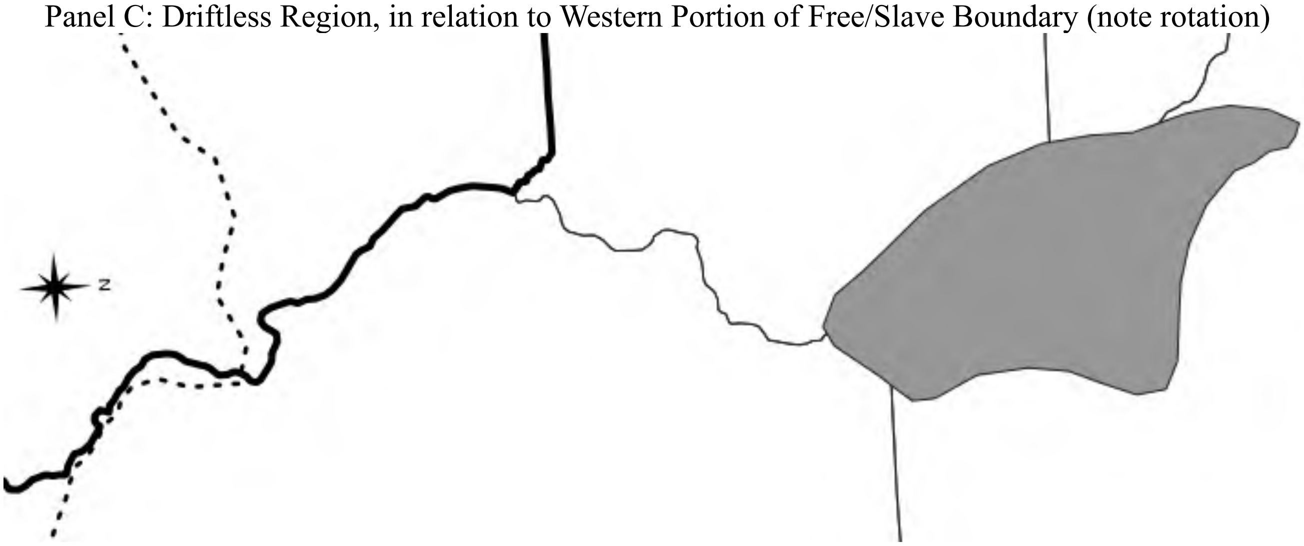
Appendix Figure C.2: Three Close-up Views of Terminal Moraine

Panel A: Crossing Points into Kentucky and Missouri

Panel B: Detail in Northern Pennsylvania and Environs



Panel C: Driftless Region, in relation to Western Portion of Free/Slave Boundary (note rotation)



Notes: see note for Appendix Figure C.1. The Driftless Region is graded in gray in Panel C.

Appendix D: Spatial Correlation

We now consider alternative strategies for assessing the precision of our estimates above, in light of the spatial correlation in the data. A county should not be considered independent of its immediate neighbors, because so many of their outcomes have determinants that are either common or highly correlated. The strategy above is to use 15 bins of longitude as clusters, which follows on the work of Bester, Conley, and Hansen (2011) as a computationally efficient procedure to account for spatial correlation, at least within the stated groups. The averages across these 15 groups themselves exhibit low spatial correlation, suggesting that the strategy is adequate to mop up the variation that is correlated across county observations. In Table 4, Panel A, we compare the estimated standard errors for a few clustering strategies. The first row contains the estimated coefficient and the second row contains the baseline standard error. The next row uses instead 10 groups of longitude as clusters, which inflates the standard error to some degree. The following row uses only five groups of longitude, for still larger standard errors. The statistical significance would be judged essentially the same under all three of the strategies, with the exception of farm value per county area, whose coefficient becomes marginally significant when using only five bins of longitude. In the next row we use states that the clustering variable. There is some justification for this inasmuch the policy under analysis (among other policies conceivably) vary at the state level. The strategy yields still larger standard errors, and results for rural population and farm value per county area are rendered marginally significant. For comparison, we also report standard errors under the assumption of independence, and these are considerably smaller than those using the large clusters.

We then turn to a parametric approach for dealing with space: the Conley (1999) estimator. This estimator uses a predetermined band of distance around each observation and estimates the correlation within it. (This is analogous to the Newey-West estimator for time series.) In the first row, we present the estimates assuming independence. A typical distance from one county's center to its neighbor's is 5 to 15 miles. Therefore, starting with a 10-mile band seemed appropriate. This allows an observation to be correlated with its immediate neighbors who, in turn, can be correlated with their immediate neighbors. As seen in the second row, the estimated standard errors hardly budge. Doubling the band to 20 miles increases the estimated standard error, although not by much. We present, in the remainder of the panel, results for bands out to 60 miles. Roughly speaking, the standard errors double over the ones based on an assumption of independence. This does not change the statistical significance of any of the coefficients. With this reassurance, we can move forward to explore other important differences between the free and slave regions.

Appendix Table D-1: A Whole Table on Standard Errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcomes (in natural logarithms):	Nonwhites per county acre	Whites per county acre	Rural population per county acre	Total farm acres per county acre	Improved acres per total farm acreage	Farm value per county area	Farm value per total farm acreage
Panel A: Various clusters							
Coefficient	1.899	-0.644	-0.511	-0.0239	-0.405	-0.582	-0.558
Cluster for 15 quantiles of latitude (baseline)	(0.485)	(0.142)	(0.157)	(0.146)	(0.140)	(0.252)	(0.177)
Cluster for 10 quantiles of latitude	(0.546)	(0.142)	(0.159)	(0.118)	(0.153)	(0.252)	(0.186)
Cluster for 5 quantiles of latitude	(0.750)	(0.189)	(0.195)	(0.121)	(0.198)	(0.336)	(0.237)
Cluster for state	(0.731)	(0.338)	(0.327)	(0.282)	(0.143)	(0.395)	(0.212)
No clustering	(0.203)	(0.104)	(0.0935)	(0.0806)	(0.0566)	(0.132)	(0.0831)
Panel B: Conley standard errors, various cutoff distances							
Independent	(0.197)	(0.098)	(0.075)	(0.064)	(0.051)	(0.115)	(0.079)
10 miles	(0.197)	(0.099)	(0.075)	(0.064)	(0.051)	(0.115)	(0.079)
20 miles	(0.211)	(0.102)	(0.077)	(0.065)	(0.055)	(0.121)	(0.085)
30 miles	(0.250)	(0.114)	(0.092)	(0.076)	(0.068)	(0.144)	(0.101)
40 miles	(0.293)	(0.127)	(0.107)	(0.087)	(0.080)	(0.168)	(0.117)
50 miles	(0.332)	(0.138)	(0.120)	(0.097)	(0.092)	(0.189)	(0.132)
60 miles	(0.369)	(0.149)	(0.132)	(0.107)	(0.103)	(0.208)	(0.145)

Notes:

Appendix E: Agricultural Activities and Wealth Distributions

E.1. Farm Output

Delving deeper into the Census of Agriculture allows us to observe crucial differences in farming operations, in crop choices and farm sizes. Figure 5 graphs the point estimates and confidence intervals for key variables related to farm production activities in 1860. We examine the top 20 farm products in the 300-mileband around the border. We will refer to these as crops although four– butter, cheese, wool, and honey—are technically animal products. Note that the slave-free border was sufficiently far north that cotton, rice, and sugar cane are not among the top 20 farm products. Also note to avoid clutter, we drop displaying results for the "donut" sample from this point forward. The extent of the production activity in each county is normalized by total farm acreage (US Census Office, 1864b).

Analysis of the crop data reveals that the small grains (wheat, rye, oats, barley, and buckwheat) are less typically common in the slave region. They were grown in both regions, but less commonly in the slave side. The gaps for rye and buckwheat were small, but that for barley very large.[1] Animal products (and their inputs such as hay and clover seed) were also typically less common in the slave region. Corn showed no difference at the border and only small differences in the widest band (in line with the standard findings of economic history that southern farms and plantations in the antebellum period were generally self-sufficient in maize.) The crops that were more common in the slave region were tobacco and hemp. Indeed, the differences for tobacco are most apparent. Tobacco had year-round labor requirements with intensive activity levels in close proximity, which facilitated direct surveillance by supervisors.
Farmers on either side of the border could grow tobacco, but those on the slave side specialized more. There was less specialization in wheat and in butter production. This may be compared with the stated intentions of Thomas Jefferson. Jefferson was one of the authors of the Northwest Ordinance, which opened up the territory north of the Ohio River to white settlement. He also promoted the idea that the country would be better off if it were populated principally by free yeoman farmers. Small-scale operators might be expected to be engaged in producing grain, dairy, and other diversified outputs. The producers on slave side had access to a technology that those on non-slave side did not—they could coerce non-family labor to join their work-force and attain a larger scale of operations with greater ease. (See Fleisig 1976, Naidu 2020).

The distribution of farm sizes also looks different on the two sides of the border. Figure 6 uses the Census of Agriculture data to provide a contrast of scale of farm operations. Operations of 500-999 acres were significantly (in both statistical and economic terms) more common in the slave region. There is a lower prevalence of farms at the yeoman's scale, in the 50-99 acreage range. The gap is nearly significant at conventional levels at the border but becomes statistically significant at the 95-percent level for the widest band considered, within 300 miles.

---

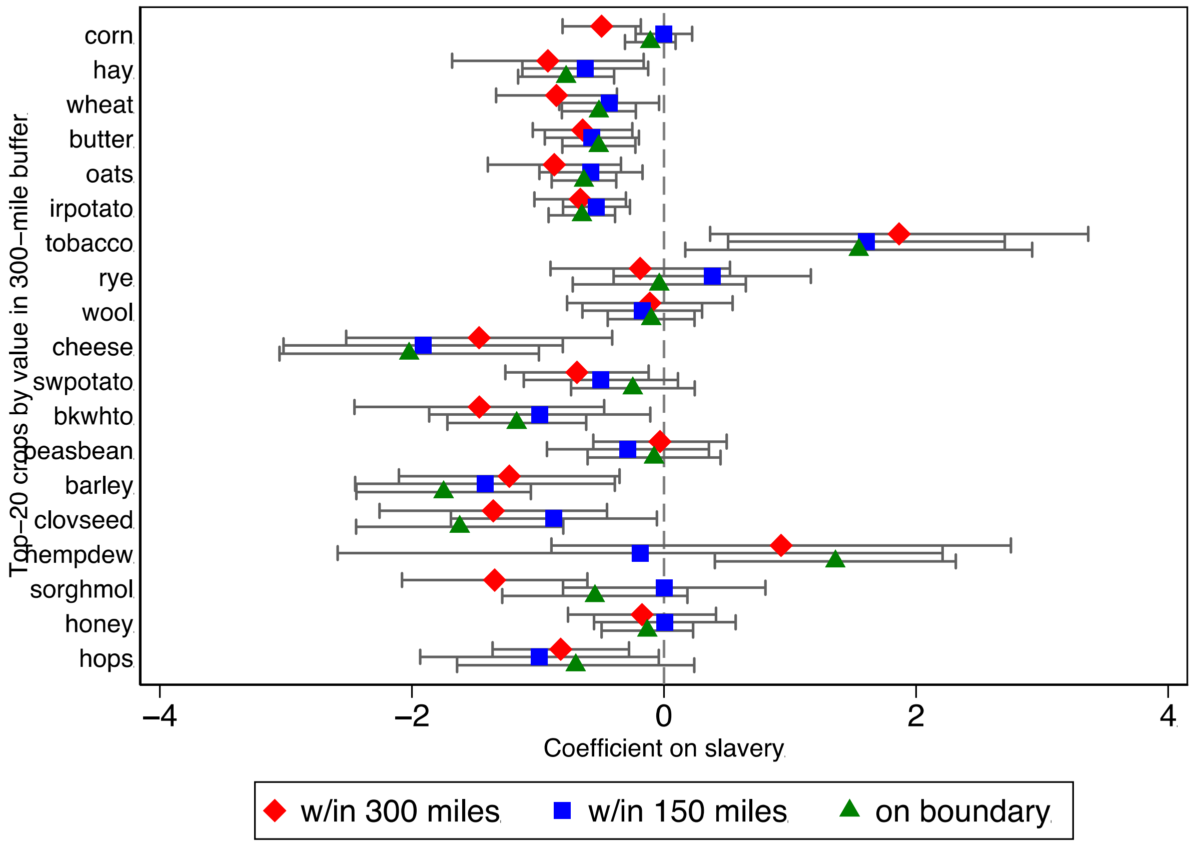
[1] To speculate, one notes rye and buckwheat are hardy, low-input crops, grown on marginal lands. Barley was a low value to weight crop used to brew beer. Hops, also low in the slave region, was also used to flavor beer. Beer was not produced with consumption by the enslaved population in mind. See Gray (1933).

E.2: Household-Level Wealth Distributions

We can extend the analysis of property holding by looking at Census micro data on wealth. In 1860, the Census asked free people about the value of the real property and personal property that they owned. This is the universe of free people. Personal property in this case includes enslaved African Americans, so there is a mechanical aspect to the change at the border.

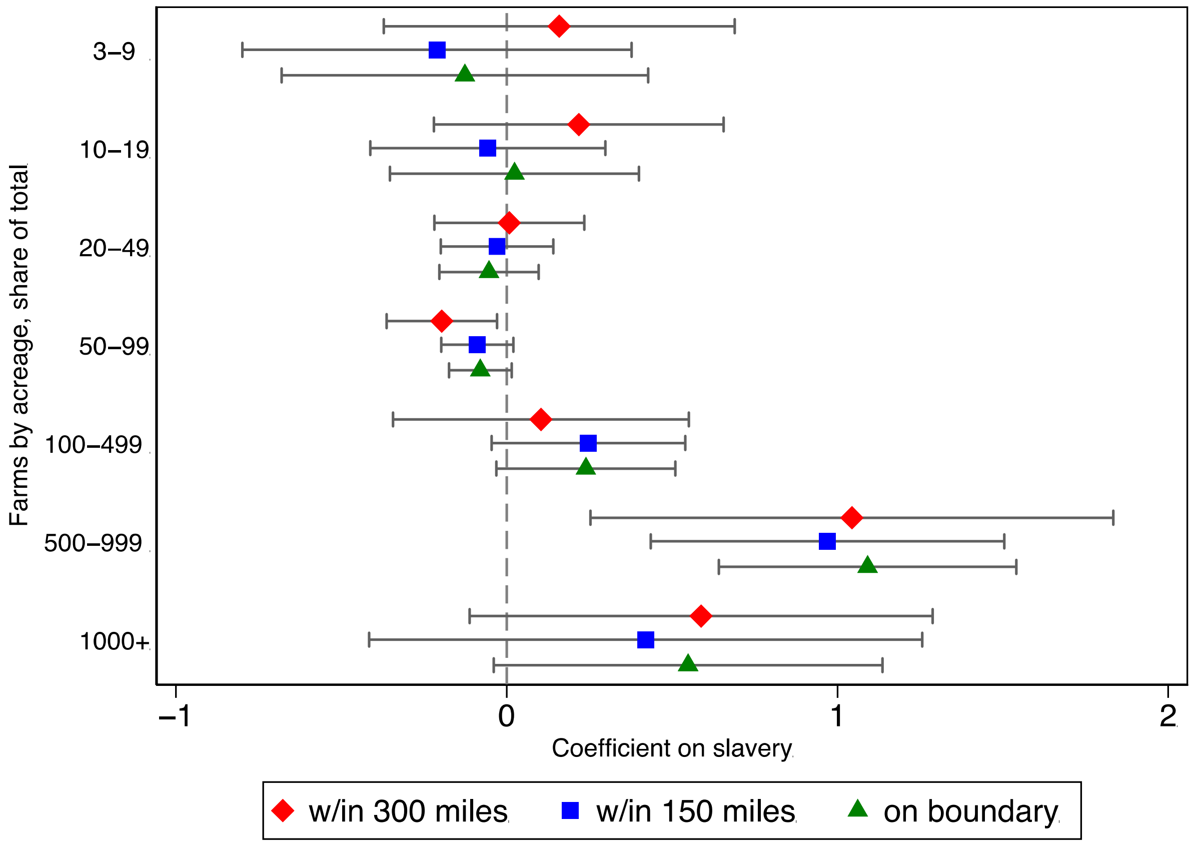
The panels of Figure 7 graph the probability density functions for wealth (measured on the log scale) in the border area in 1860. The panels do so for real property, personal property, and total property, by household. (Results are similar if we restrict to the rural areas.) The Figure reveals extra mass at the high levels of wealth on the slave side, both in personal property and in real property. The Figure also reveals missing mass starting around $2,000, which seems to be shifted out to above $10,000. Again, the small-scale property holders appear less prevalent in the slave region.

Appendix Figure E-1: Crops



Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero.

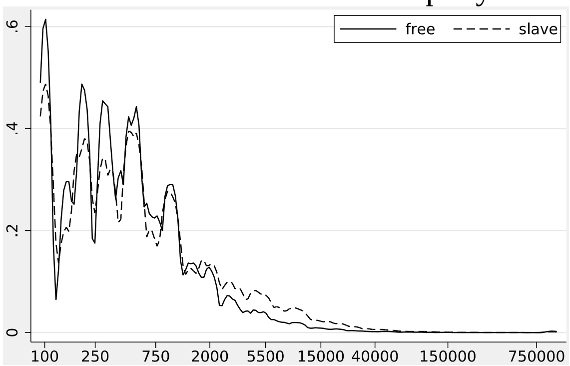
Appendix Figure E-2: Farm sizes



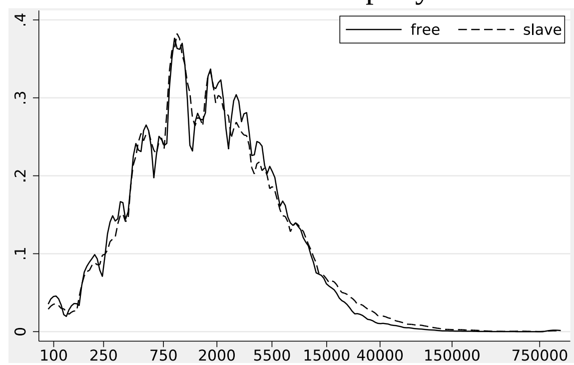
Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero.

Appendix Figure E-3: Wealth at the household level

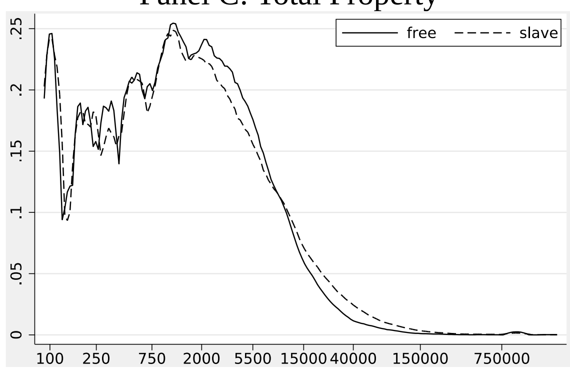
Panel A: Personal Property



Panel B: Real Property



Panel C: Total Property



Notes: This figure presents estimates of the probability density function for log wealth at the household level in the border counties in 1860. As distinct from preceding figures, these are estimated using the microdata from the full-count 1860 census.

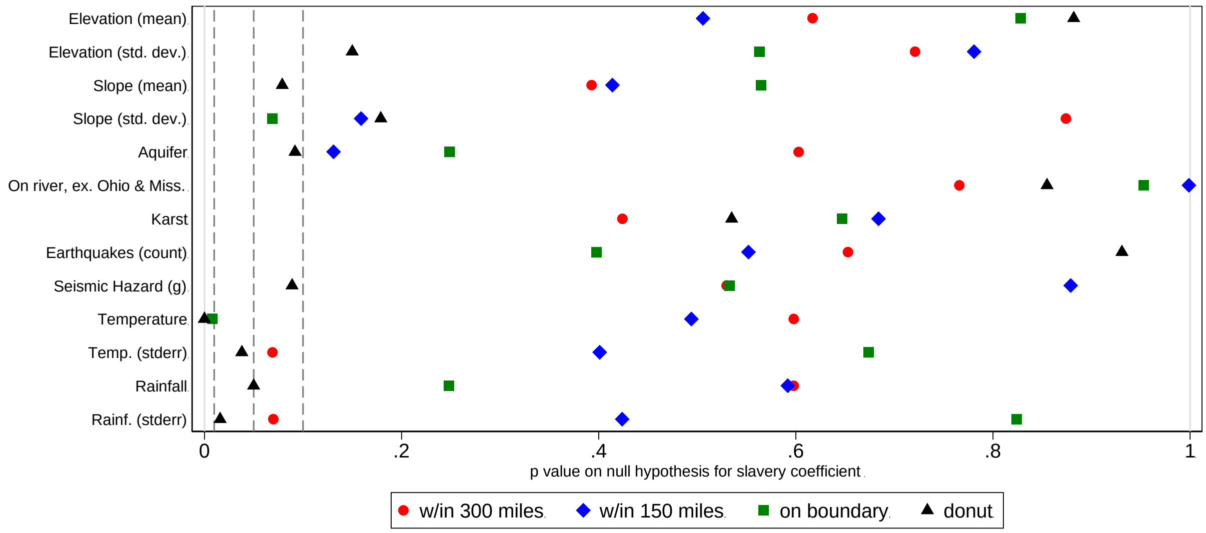
Appendix F: Additional Sets of Results (Sensitivity Analysis for Main Results)

Figures:

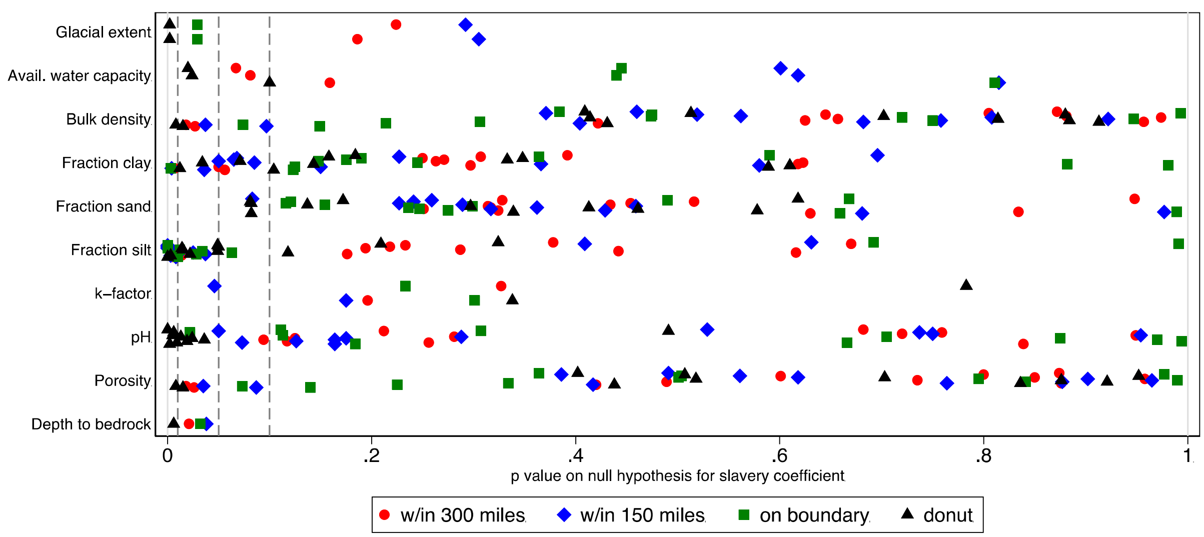
Tables:

Appendix Figure F-0: Environmental Factors

Panel A: Miscellaneous

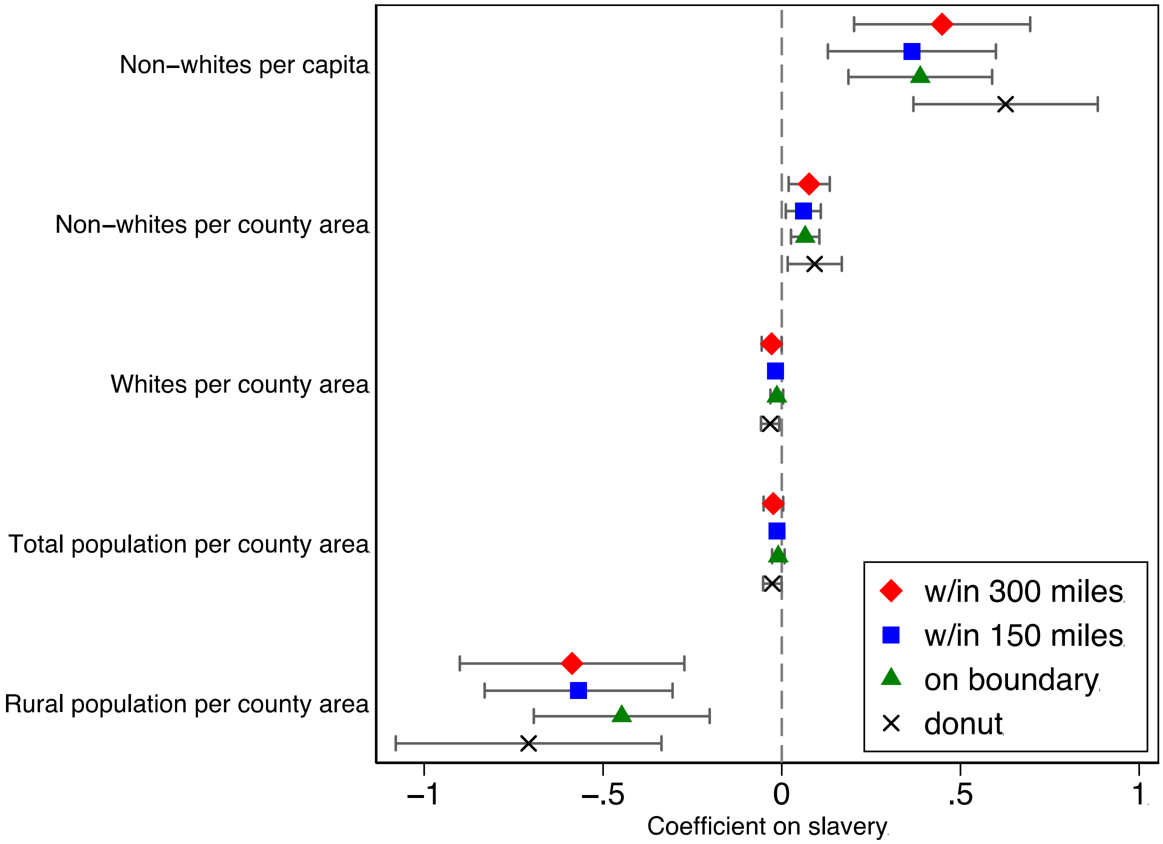


Panel B: Soil Variables (by group)



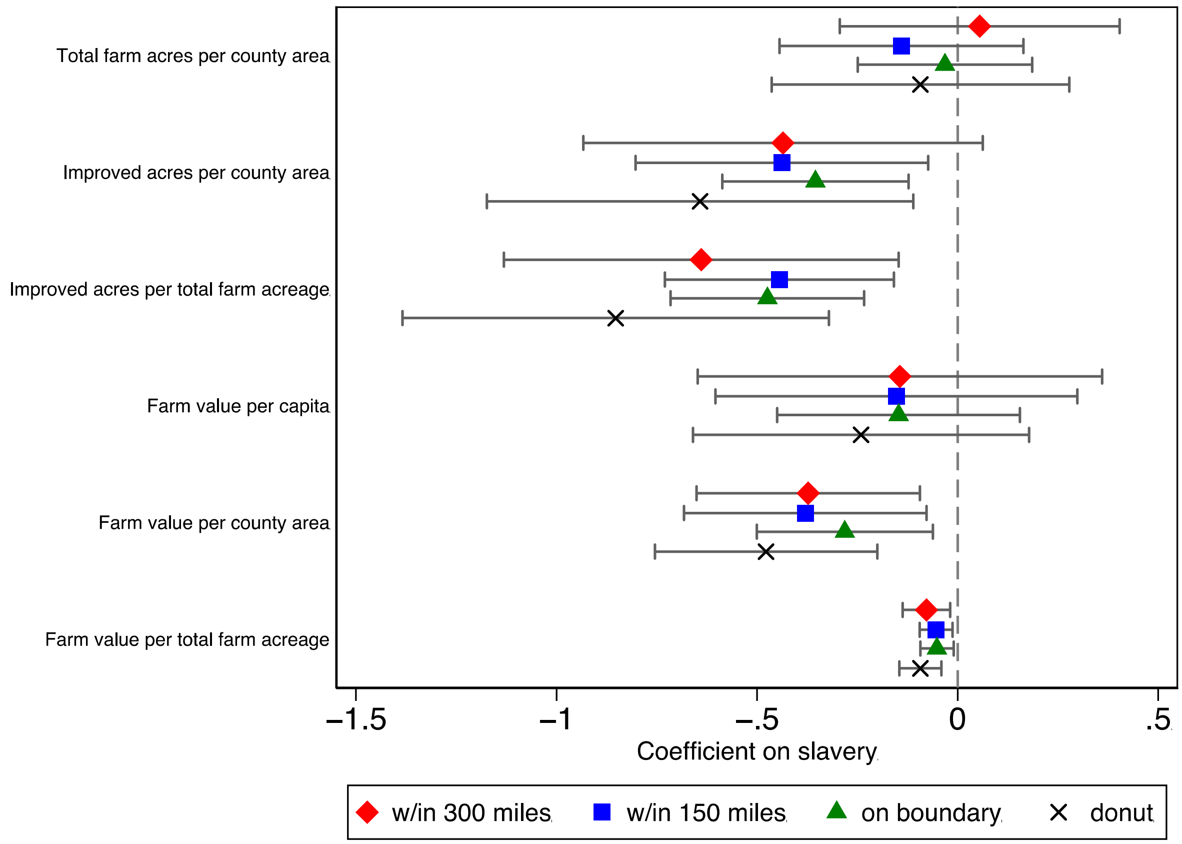
Notes: This figure presents probability values for the null hypotheses that the coefficient on slavery equals zero. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol is associated with the test of the null hypothesis for the outcome indicated in the row label and for various samples of counties. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed lines denote standard cutoffs at 1%, 5%, and 10%. See the text for variable sources and definitions.

Appendix Figure F-1: Population (z-score)



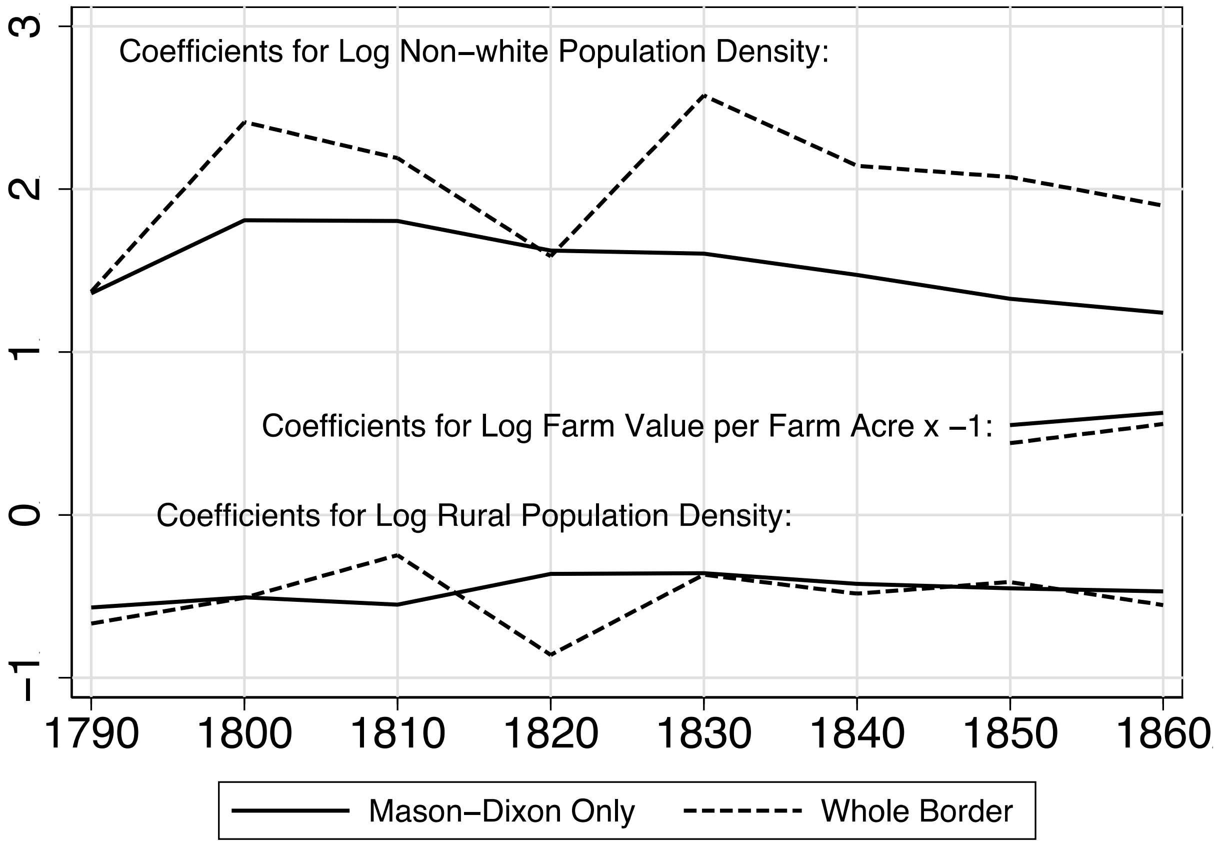
Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero. The outcomes are transformed into z-scores.

Appendix Figure F-2: Land Use and Land Value (z-score)



Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero. The outcomes are transformed into z-scores.

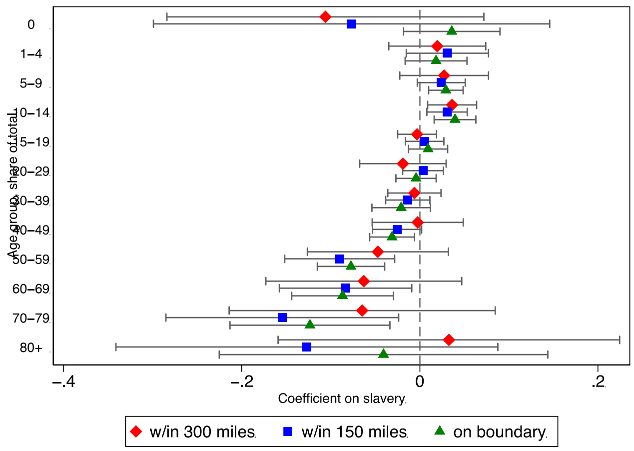
Appendix Figure F-3: Effects on Nonwhite and Rural Population, Various Years



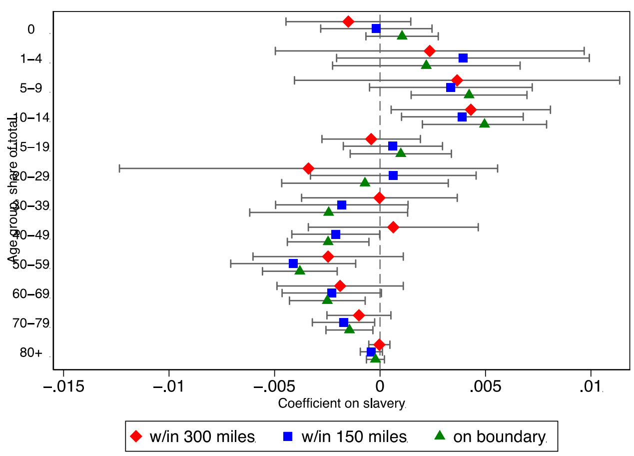
Notes: This figure presents point estimates for the coefficient on slavery for the outcomes indicated in the graph and for various samples of counties. The regression specification is the default: polynomial controls for longitude and for distance to the border, clustered errors by 15 bins of longitude, and weights according to land area. The base sample consists of counties in the 300-mile buffer. The "whole border" sample uses all of the available counties in each year. The Mason-Dixon sample uses only those counties whose closest free-slave border abuts Pennsylvania or New Jersey. (All of the land in the 300-mile buffer is covered by a county by 1810. All of the land in the Mason-Dixon sample is covered by counties for 1790 forward.)

Appendix Figure F-4: Age Composition

Panel A: Logarithm



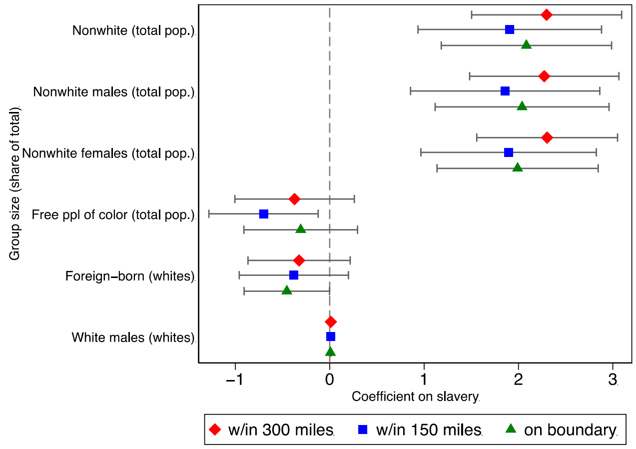
Panel B: Inverse Hyperbolic Sin



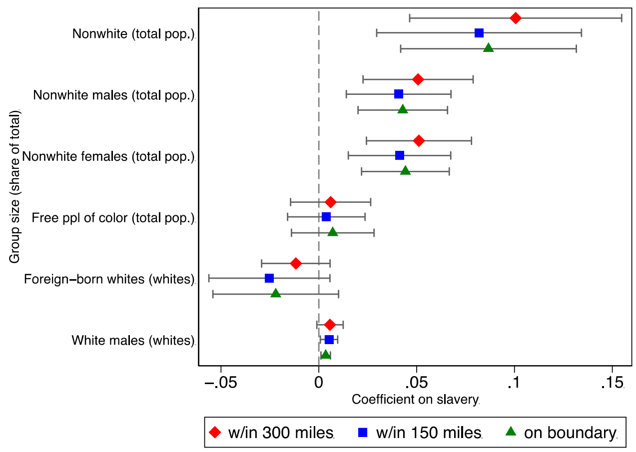
Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero. For the top panel, the outcomes are transformed into natural logarithms. For the bottom panel, the outcomes are transformed with the inverse hyperbolic sin (asinh).

Appendix Figure F-5: Race and Gender

Panel A: Logarithm

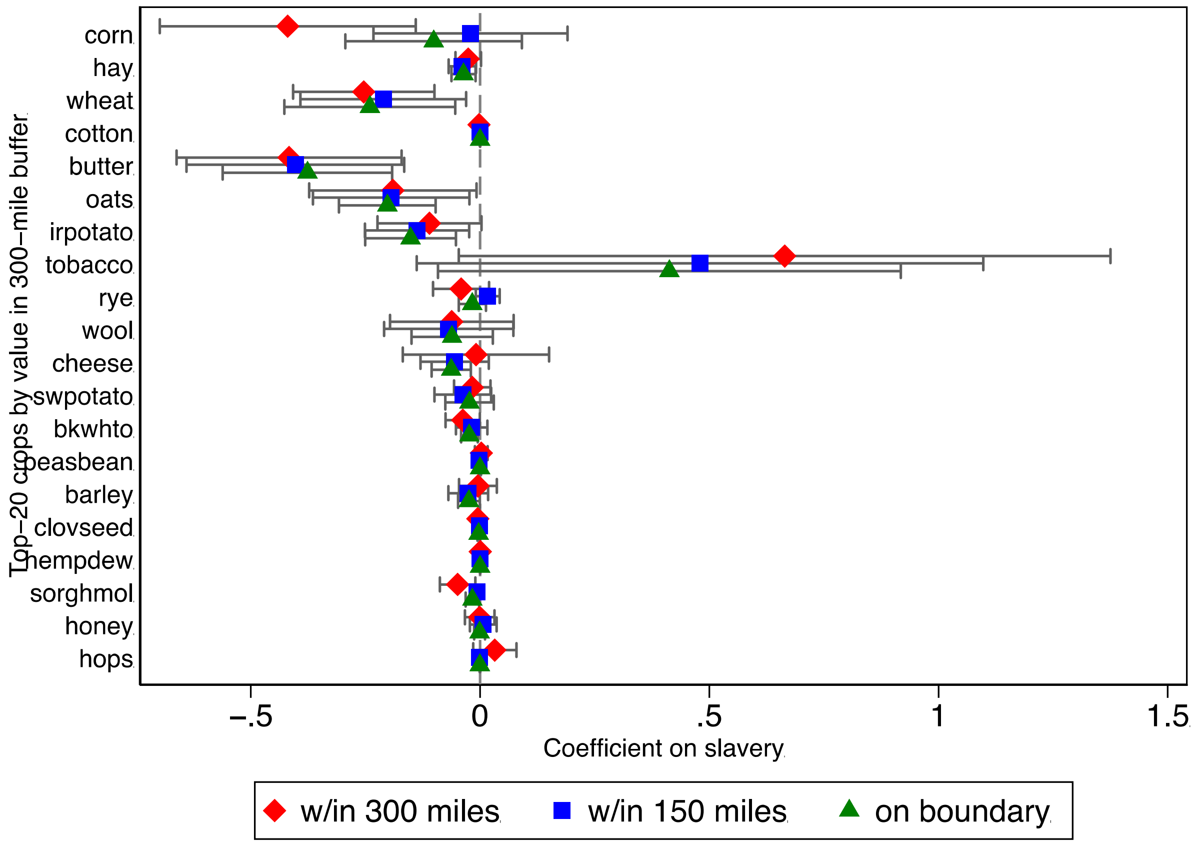


Panel B: Inverse Hyperbolic Sin



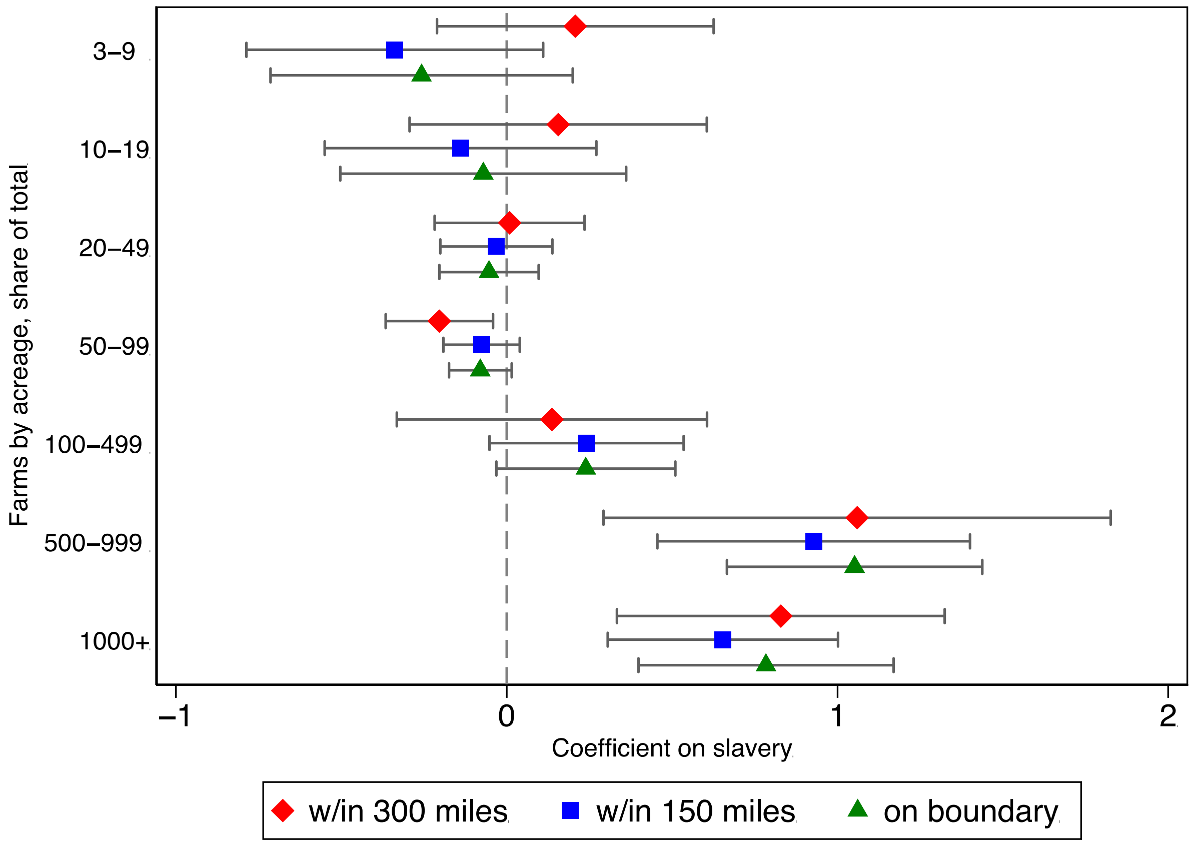
Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero. For the top panel, the outcomes are transformed into natural logarithms. For the bottom panel, the outcomes are transformed with the inverse hyperbolic sin (asinh).

Appendix Figure F-6: Crops (asinh transform)



Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero. The outcomes are transformed with the inverse hyperbolic sin (asinh).

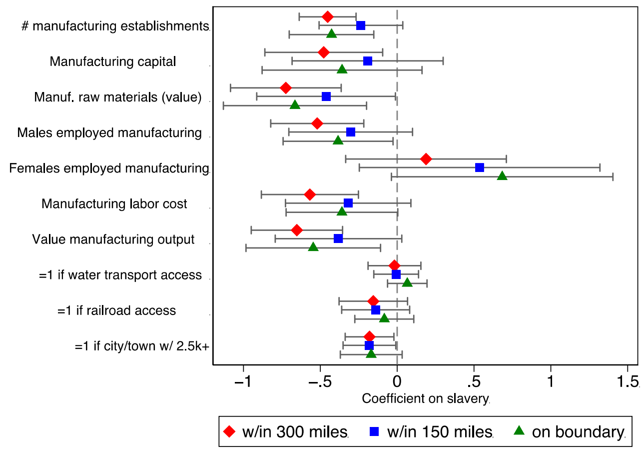
Appendix Figure F-7: Farm sizes (asinh transform)



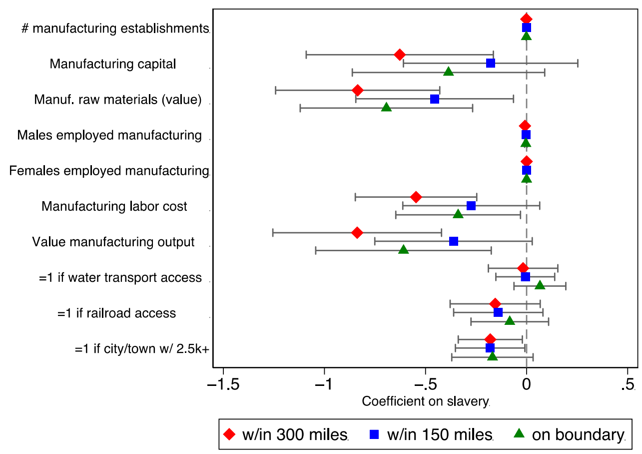
Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero. The outcomes are transformed with the inverse hyperbolic sin (asinh).

Appendix Figure F-8: Structural Transformation

Panel A: Logarithm

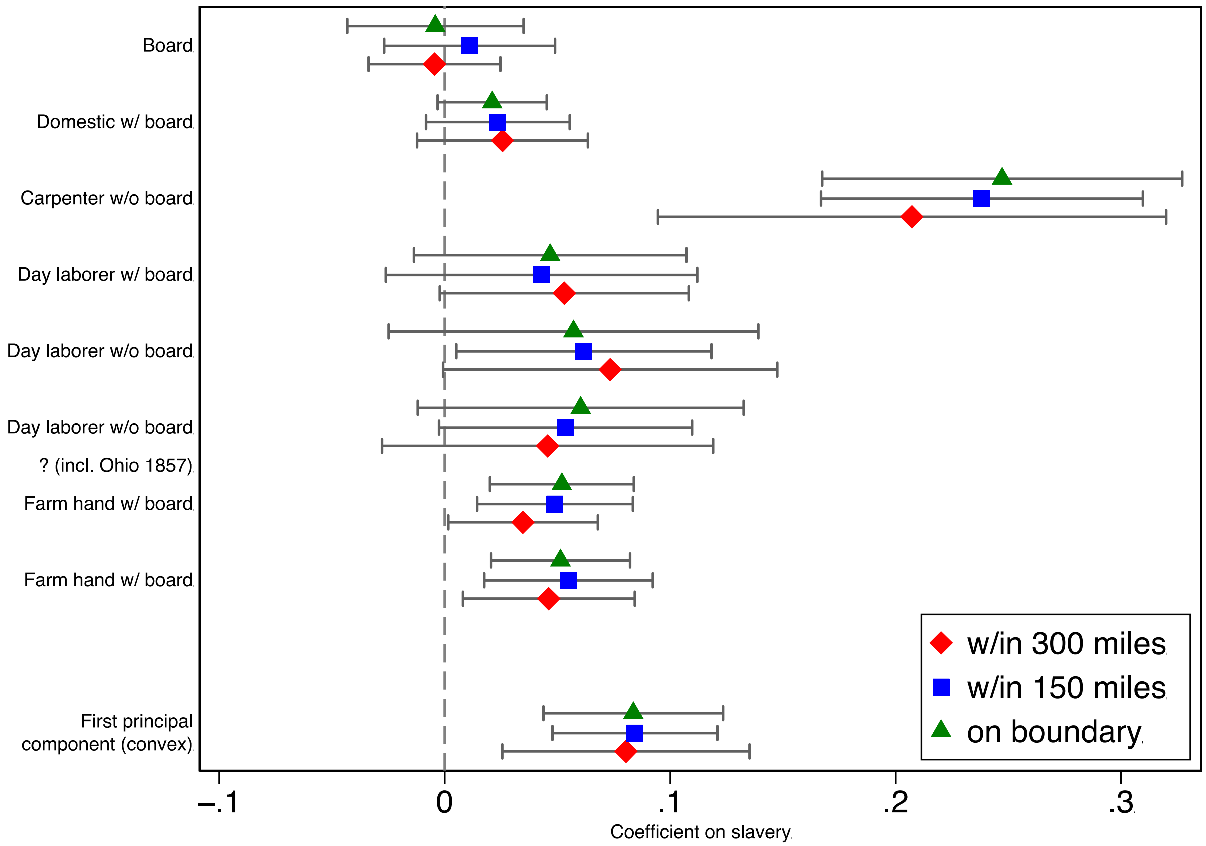


Panel B: Inverse Hyperbolic Sin



Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero. For the top panel, the outcomes are transformed into natural logarithms. For the bottom panel, the outcomes are transformed with the inverse hyperbolic sin (asinh).

Appendix Figure F-9: Wages (levels)



Notes: This figure presents point estimates and confidence intervals for the coefficient on slavery for the outcomes indicated in the row label and for various samples of counties. Point estimates are denoted with symbols within horizontal bands denoting 95-percent-confidence intervals. Standard errors are estimated using 15 quantiles of longitude as clusters. Each symbol type the notes a distinct sample: red diamond for counties within 300 miles of the boundary, blue square for counties within 150 miles of the boundary, and green diamond for counties adjacent to the boundary. The vertical, dashed line denotes a null hypothesis of zero.

Appendix Table F-1. Summary statistics, select variables and samples.

Variable	Units	Whole Obs.	Whole Mean	Whole Std. Dev.	Free side Obs.	Free side Mean	Free side Std. Dev.	Slave side Obs.	Slave side Mean	Slave side Std. Dev.	Difference in mean
Panel A: Counties within 300 miles of boundary											
Nonwhite Population,	Per Thous.	[1362]	5.3	(17.7)	[659]	0.9	(8.7)	[703]	9.6	(22.5)	-8.7
per County Acre	Logs	[1280]	-7.0	(2.4)	[579]	-8.6	(2.0)	[701]	-5.6	(1.7)	-3.0
White Population,	Per Thous.	[1362]	46.9	(342.1)	[659]	69.0	(467.0)	[703]	25.2	(130.3)	43.9
per County Acre	Logs	[1362]	-3.7	(1.2)	[659]	-3.4	(1.5)	[703]	-4.0	(0.7)	0.57
Rural Population,	Per Thous.	[1362]	41.1	(30.9)	[659]	51.1	(37.1)	[703]	31.2	(18.8)	19.8
per County Acre	Logs	[1357]	-3.6	(1.1)	[656]	-3.5	(1.5)	[701]	-3.7	(0.7)	0.18
Total Farm Acreage,	Levels	[1356]	0.59	(0.29)	[656]	0.55	(0.29)	[700]	0.63	(0.29)	-0.07
per County Acre	Logs	[1356]	-0.79	(1.00)	[656]	-0.97	(1.26)	[700]	-0.61	(0.61)	-0.36
Improved Acreage,	Levels	[1356]	0.41	(0.20)	[656]	0.51	(0.20)	[700]	0.32	(0.15)	0.19
per Farm Acre	Logs	[1356]	-1.03	(0.57)	[656]	-0.78	(0.51)	[700]	-1.27	(0.51)	0.49
Farm value ($),	Levels	[1356]	11.57	(14.01)	[656]	15.71	(17.19)	[700]	7.51	(8.12)	8.20
per County Acre	Logs	[1356]	1.75	(1.48)	[656]	1.95	(1.81)	[700]	1.54	(1.03)	0.41
Farm value ($),	Levels	[1356]	17.50	(37.36)	[656]	24.05	(51.43)	[700]	11.07	(9.27)	12.98
per Farm Acre	Logs	[1356]	2.53	(0.81)	[656]	2.92	(0.72)	[700]	2.15	(0.70)	0.77
Panel B: Counties on the boundary											
Nonwhite Population,	Per Thous.	[142]	4.2	(6.6)	[71]	1.9	(3.1)	[71]	6.8	(8.4)	-5.0
per County Acre	Logs	[140]	-6.9	(2.1)	[69]	-7.6	(2.1)	[71]	-6.0	(1.8)	-1.6
White Population,	Per Thous.	[142]	67.3	(83.4)	[71]	77.3	(88.3)	[71]	56.0	(76.4)	21.3
per County Acre	Logs	[142]	-3.0	(0.7)	[71]	-2.8	(0.7)	[71]	-3.2	(0.7)	0.42
Rural Population,	Per Thous.	[142]	55.8	(33.4)	[71]	64.4	(36.3)	[71]	45.9	(26.8)	18.6
per County Acre	Logs	[142]	-3.1	(0.6)	[71]	-2.9	(0.6)	[71]	-3.2	(0.6)	0.33
Total Farm Acreage,	Levels	[142]	0.65	(0.19)	[71]	0.66	(0.20)	[71]	0.64	(0.18)	0.02
per County Acre	Logs	[142]	-0.49	(0.39)	[71]	-0.48	(0.38)	[71]	-0.51	(0.40)	0.03
Improved Acreage,	Levels	[142]	0.47	(0.17)	[71]	0.52	(0.16)	[71]	0.41	(0.17)	0.11
per Farm Acre	Logs	[142]	-0.83	(0.39)	[71]	-0.70	(0.34)	[71]	-0.97	(0.40)	0.26
Farm value ($),	Levels	[142]	17.22	(18.02)	[71]	20.41	(20.76)	[71]	13.56	(13.52)	6.85
per County Acre	Logs	[142]	2.42	(0.93)	[71]	2.61	(0.93)	[71]	2.21	(0.88)	0.40
Farm value ($),	Levels	[142]	23.89	(20.38)	[71]	27.68	(22.41)	[71]	19.56	(16.95)	8.11
per Farm Acre	Logs	[142]	2.91	(0.69)	[71]	3.09	(0.66)	[71]	2.72	(0.67)	0.37

Notes: