Firm Heterogeneity, Technological Adoption, and Urbanization: Theory and Measurement

Alex W. Chernoff*

May 3, 2017

Abstract

This paper develops a model of firm heterogeneity, technological adoption, and urbanization. This theoretical framework is developed with the objectives of estimating the effect of mechanical steam power on urbanization and welfare in the nineteenth century. In the context of the theoretical model, urbanization is measured by population density, and welfare is measured by household real income. I derive statistics that measure the effect of a new technology on township urbanization and aggregate welfare. The empirical application of the paper estimates these statistics using nineteenth century firm level data on mechanical steam power in the Canadian manufacturing sector, and township level population data. The results indicate that the introduction of steam power increased welfare by 5.8 percent. To evaluate the contribution of steam power to urbanization, I compare observed township population density growth over the period 1861-71, to the model predicted change in township population density from the introduction of steam. The two series are positively correlated, with the modeled change accounting for 6.0 percent of the observed variation. This result suggests that population re-sorting in response to the introduction of mechanical steam power had a positive, but economically modest, effect on urbanization in the mid-nineteenth century.

JEL classification: O14, R13, N61
Keywords: Welfare, Firm Heterogeneity, Urbanization, Industrialization, Steam

Preliminary and Incomplete. Please do not cite without author's permission.

*Bank of Canada, email: achernoff@bankofcanada.ca. The views in this paper are those of the author and do not necessarily reflect those of the Bank of Canada. This paper is a substantially revised version of my job market paper. I gratefully acknowledge helpful feedback provided by Brant Abbott, Geoffrey Dunbar, Beverly Lapham, Frank Lewis, Huw Lloyd-Ellis, Ian Keay, Oleksiy Kryvtsov, Marvin McInnis and Diego Restuccia. This paper benefited from discussions with participants at the Economic History Association Meetings and Cliometric Society Conference in 2015, and the 2014 Canadian Network for Economic History Conference. I acknowledge the helpful comments of seminar participants at Université Laval, Wilfrid Laurier University, UC Davis, Monash University, University of Saskatchewan, and the University of Alberta.

1 Introduction

Technological adoption and urbanization are defining features of the industrial revolution, however there is debate in the literature over the interaction between these two phenomena. For example, the empirical work of Rosenberg and Trajtenberg (2004) and Kim (2005) reach contradictory conclusions regarding the importance of steam power in fostering urbanization in the U.S. during the late nineteenth century. To advance this debate, this paper develops and estimates a model of firm heterogeneity, technological adoption, and urbanization. The empirical application of the paper considers the use of mechanical steam power in Canadian manufacturing in the nineteenth century. Using a theoretical framework and empirical analysis that incorporates firm heterogeneity and economic geography, this paper estimates the effect of steam power on welfare and urbanization.

The theoretical section of the paper embeds a model of firm heterogeneity and technological adoption into the new economic geography framework of Helpman (1998). In the model, firms make a discrete choice to use one of several different technologies that vary in terms of productivity. Each mode of production has a technology and township specific fixed cost of adoption. Labor moves freely between townships that also differ in their endowment of residential land. Households gravitate toward high wage townships where firms readily adopt the frontier technology. However, the fixed supply of land in each township creates a congestion externality that lowers welfare in high wage townships. This tension is resolved by labor mobility, which equalizes household utility between densely populated high wage townships, and sparsely populated low wage townships.

The model is developed to achieve the two main objectives of this paper. The first objective is to estimate a theoretically derived statistic that measures the welfare gains from the introduction of mechanical steam power. In the context of the model, welfare gains are equal to the change in household real income relative to a counterfactual economy where steam power is unavailable. The second objective is to derive a statistic that measures the contribution of mechanical steam power to urbanization. For the purposes of this paper, urbanization is defined as the change in population density at the township level. It is important to note that the model predicted effects of steam power on welfare and urbanization relate to a *static* comparison of the observed and counterfactual economy. In particular, in the counterfactual economy the land endowment of each township and the aggregate population are held constant at their levels in the observed economy. Therefore, the modeled changes in township population densities occur as a result of population re-sorting in response to the introduction of steam power.

In the empirical application, I estimate the model using nineteenth century firm level data on steam power adoption in the Canadian manufacturing sector. These data offers a number of advantages for studying the effects of steam power on welfare and urbanization. In particular, it is possible to study population density at a highly disaggregated spatial resolution by using Geographic Information Systems (GIS) township maps for mid-nineteenth century Ontario. A second advantage of the data is their comprehensive nature. My analysis uses firm level data on *all* firms enumerated in the 1871 Census of Manufacturing for the province of Ontario. Ontario accounted for over half of Canadian manufacturing output, and prior research shows that manufacturers in the province were broadly comparable to manufacturing firms in the Northeastern U.S. during this era (Inwood and Keay, 2005).

An empirical challenge in estimating the model relates to identifying the effect of steam power on firm level productivity. Using instrumental variables estimation, I exploit terrain slope as an exogenous source of geographic variation. The idea for the identification strategy comes from Bishop and Muñoz-Salinas (2013), who show that slope is an important factor in explaining the location of historic watermills in England and Scotland. The authors note that upstream dams were required to create reservoirs that were used to power watermills. In steeper channels of a river, the height and cost of building upstream dams was lower, as was the cost of the infrastructure needed to bring water from the reservoir to the waterwheel. In the context of the model, terrain slope lowers the fixed cost of using water power, and *decreases* the adoption rate of steam power. For my instrumental variable, I estimate the average slope within each township using GIS analysis and data sourced from the NASA Shuttle Radar Topography Mission. The IV estimation results indicate that steam power adoption increased firm level labor productivity by 21.9 percent relative to water power. The welfare statistic is estimated using the IV regression results and the estimates of a second regression equation that is derived from the theoretical model. The results indicate that the introduction of steam power increased welfare by 5.83 percent, as measured by the change in household real income.

To study the relationship between steam power adoption and urbanization, I use the estimated model to predict the population densities of the 440 Canadian townships in the counterfactual economy where steam power is unavailable. The model predicted change in township population densities quantifies the extent of re-sorting that occurs with the introduction of steam power. To estimate the contribution of steam power to urbanization through this channel, I compare the model predicted versus actual growth in population density for Canadian townships in the mid-nineteenth century. The correlation between the two series is positive and highly significant, however the model predicted change in population density only accounts for 6.0 percent of the observed variation population density growth. This result suggests that the population re-sorting induced by mechanical steam power adoption had a positive, but economically modest, effect on urbanization in the mid-nineteenth century.

Theoretically, this paper contributes to the literature on firm heterogeneity, technological change, welfare, and urbanization. In the international trade literature, papers by Yeaple (2005), Lileeva and Trefler (2010), and Bustos (2011) have established that trade liberalization can induce technological upgrading through increased market access. This paper draws on the theoretical frameworks used in this literature to study the effect of a new technology on urbanization. In particular, my theoretical model extends Bustos (2011), by incorporating multiple technologies, free mobility of labor, and regional resource endowments. With respect to geography, my theoretical framework is similar to Krugman (1995a), Krugman (1995b), Helpman (1998), Michaels et al. (2012), and Redding (2012). I contribute to this literature by embedding a model of firm heterogeneity and technological adoption into the economic geography model of Helpman (1998), and by illustrating how this theoretical framework can be used to estimate the effects of a new technology on welfare and urbanization.

Empirically, this paper contributes to the literature on steam power, industrialization, and urbanization in the nineteenth century.[1] To my knowledge, this is the first paper to estimate the effect of steam power

[1]While the literature on steam power and urbanization is sparse, there is an extensive literature documenting the importance of steam power to growth and industrialization during this era. For the U.S., see Temin (1966), Atack (1979), Atack et al. (1980), Hunter (1985) and Atack et al. (2008). For Canada, see Bloomfield and Bloomfield (1989) and Inwood and Keay (2012). For the U.K., see Crafts (2004).

on urbanization, using a theoretically derived empirical framework that explicitly models the endogenous relationship between technological adoption and urbanization. Rosenberg and Trajtenberg (2004) find that the diffusion of the Corliss steam engine led to an increase in urbanization in North and Mid-Atlantic US states during the late nineteenth century. The authors argue that steam power diffusion increased urbanization by freeing manufacturers from the locational constraints associated with water power. Rosenberg and Trajtenberg (2004) acknowledge the importance of the endogenous relationship between technological adoption and agglomeration, which is not incorporated in their empirical analysis. Rosenberg and Trajtenberg (2004) conclusions contrast with Kim (2005), who argues that steam power diffusion was not a primary factor in explaining urbanization in the U.S. during the late nineteenth century. Kim (2005) uses firm level data from the decennial U.S. Manufacturing Censuses and finds that the diffusion of steam power accounted for no more than an 8-10 percent increase in urbanization between 1850-1880. Kim (2005) acknowledges that general equilibrium effects confound his empirical analysis, noting that his estimates are likely biased upward due the endogenous relationship between steam power adoption and urbanization. The conclusions of Rosenberg and Trajtenberg (2004) and Kim (2005) stress the importance of general equilibrium effects in explaining the relationship between steam power adoption and urbanization, however neither paper analyzes these effects. This paper fills this void in the literature by estimating the effect of steam power adoption on urbanization, using an empirical framework that explicitly addresses the endogenous relationship between technological adoption and urbanization.

The remainder of the paper is organized as follows: section 2 describes the data that are used in this paper, section 3 presents the theoretical model, section 4 presents the empirical application, and section 5 concludes.

2 Firm Heterogeneity, Steam Power, and Urbanization

This section introduces the data that are used in the empirical application, and presents a number of empirical regularities that guide the development of the theoretical framework of the paper.

Firm level data from the Canadian Industry in 1871 Project (CANIND71) are used in this paper. The CANIND71 dataset is a digitized record of all manufacturing establishments enumerated in the 1871 Manufacturing Census of Canada. The empirical analysis uses all CANIND71 firm level records for the province of Ontario. The data from Ontario are selected for the analysis based on the availability of GIS maps for Ontario townships in 1871, which enable analysis of population density at a very high level of spatial resolution. Inwood and Keay (2005) find that Ontario manufacturers were smaller and slightly less productive, but otherwise comparable to manufacturing firms in the North-Eastern U.S. during this era. Ontario had the largest share of Canadian manufacturing output in 1871, accounting for over half of the aggregate value of goods produced.[2] The Province is centrally located and was already well integrated with the rest of North America by water and rail networks in 1871.

Similar firm level data are available for other countries, notably including the Atack and Bateman (1999) samples from the nineteenth century U.S. Manufacturing Censuses. However, it is difficult to study the effect of steam power on urbanization in the U.S., due to the lack of township or municipal level data

[2] 1871 Census of Canada, Volume III, p. 455.

during the mid-nineteenth century. The township level data used in this paper are comparable in spatial aggregation to U.S. minor civil divisions (MCD), which are used in Michaels et al. (2012). To the best of my knowledge, MCD boundary files are only available from 1880 onwards, and thus fail to capture the major era of steam power diffusion in the U.S. during the mid-nineteenth century.[3]

The comprehensive nature of the CANIND71 database is another reason why the data are well suited to the analysis of this paper.[4] Having a record of *all* manufacturing establishments ensures that the estimated welfare and urbanization statistics are representative. Furthermore, the large sample size also ensures that the the asymptotic properties of the econometric analysis can be reliably invoked.

Table 1: Production Profile by Motive of Power

	Non-Mechanized	Water	Steam	Full Sample
Employees	3.041	7.091	14.76	5.020
	(0.0625)	(0.406)	(0.713)	(0.117)
Fixed Capital	664.1	5,475	8,082	2,210
	(19.04)	(364.5)	(482.4)	(80.55)
Revenue	2,625	15,364	23,010	6,811
	(80.72)	(864.2)	(1,422)	(225.2)
Value Added	1,300	4,430	11,362	2,955
	(29.59)	(288.5)	(892.7)	(121.7)
Observations	12,088	2,114	2,008	16,210

Standard errors for the mean estimates are in parentheses. Fewer than one percent of firms reported using multiple sources of power. For the purposes of the analysis, these firms are classified by the 'most advanced' technology used. For example, firms using non-mechanized and water power are classified as water powered. Similarly, firms using water and steam are classified as steam powered. Fixed capital, revenue, and value added are in nominal 1871 Canadian dollars.

Table 1 summarizes firms' production characteristics across the different motives of power.[5] 12.4 percent

[3] Atack et al. (2008) show that the use of steam power in U.S. manufacturing grew fastest during the period 1850-1870. In the late nineteenth and early twentieth century steam was gradually overtaken by electricity as the dominant source of horsepower in U.S. manufacturing, (Jovanovic and Rousseau, 2005). At a higher level of spatial aggregation, digitized boundary files are available for the U.S. counties in the mid-nineteenth century. However, Michaels et al. (2012) note that U.S. counties often combine urban and rural areas in a single county. Therefore, U.S. county level data are not particularly well suited to studying urbanization.

[4] This contrasts with the Atack and Bateman (1999) dataset, which is based on random samples drawn from the surviving manuscripts of the decennial U.S. Manufacturing Censuses, 1850-1880.

[5] The history of the CANIND71 database is documented in Bloomfield and Bloomfield (1989). Inwood (1995) notes that multi-product firms were occasionally decomposed by enumerators into multiple manuscript entries based on the distinct industrial activities carried out by the firm. I follow Inwood's (1995) approach to identifying and reconstituting multi-product establishments. Inwood (1995, p. 364) uses the following reconstitution strategy: "The criteria for combining two or more entries into a larger and more complex firm are that they share a proprietor name within an enumeration division, that they appear immediately adjacent to each other in the manuscript schedule, and that an examination of the personal schedule 1 for the immediate area does not reveal the presence of two potential proprietors with the same name." I follow the same strategy as Inwood (1995), except that I omit the final criterion as I do not have access to the micro-data from schedule 1 of the 1871 Census of Canada. For the purposes of the analysis, I exclude firms from a number of industries that were included in the 1871 Census of Manufacturing, but would not be categorized as manufacturing industries under a contemporary classification. The excluded industries include the following Standard Industrial Classification (SIC) major industry classes: agricultural services, forestry, mining, construction, gas and water utilities, personal and business services, and trade (including repair services). I also exclude firms from the northern most Ontario townships, which were enumerated over much larger geographic areas than townships in the south.

of Ontario manufacturing firms had adopted steam power in 1871. By comparison, Atack et al. (2008) estimate that 20.1 percent of U.S. manufacturing establishments used steam power in 1870. However, this comparison is somewhat misleading as differences in the enumeration practices between the Canadian and the U.S. censuses overstate the magnitude of this difference. The U.S. census did not enumerate establishments with revenue less than $500, whereas Canadian enumerators were instructed to enumerate all establishments regardless of size, (Inwood and Keay, 2008). When the Canadian sample is censored to exclude firms falling below the U.S. census revenue threshold, the fraction of firms using steam power increases to 15.6 percent.[6]

The most striking pattern in Table 1 is the positive correlation between firm size and the use of steam power. Water powered firms were much bigger than non-mechanical firms, however firms using steam were even larger. Steam powered firms had over twice the level of employment and value added as water powered firms, and roughly fifty percent higher levels of fixed capital and revenue. This ordering across the three main technologies is replicated in the theoretical model developed in Section 3. These findings are also consistent with Atack et al. (2008), who find a positive relationship between firm size and steam power adoption in the U.S. manufacturing sector during this era.

Table 2: Descriptive Regression Analysis of Labor Productivity on Motive of Power

Variable	(1)	(2)	(3)
Steam Power	0.497***	0.444***	0.0988***
	(0.0272)	(0.0289)	(0.0271)
Water Power	0.499***	0.442***	0.0228
	(0.0340)	(0.0294)	(0.0314)
Five or Fewer Employees		-0.0434**	-0.0831***
		(0.0184)	(0.0179)
Male Employees/Employees		0.629***	0.492***
		(0.0541)	(0.0415)
ln(Population Density)		0.0977***	0.0892***
		(0.0215)	(0.0195)
ln(Capital/Labor)			0.236***
			(0.0158)
SIC Industry Controls	Yes	Yes	Yes
Observations	16,763	16,763	15,982
R-squared	0.199	0.239	0.318
p-value: Steam=Water	0.953	0.918	0.00354

Standard errors are clustered at the township level. *** p<0.01, ** p<0.05, * p<0.1. Coefficient estimates for industry indicator variables are not reported but are available upon request.

The next set of empirical results in this section explores the relationship between technological adoption and firm level productivity. Table 2 presents results from several ordinary least squares (OLS) descriptive regressions of labor productivity on steam and water power adoption. The dependent variable is the natural

[6] For the purposes of converting the U.S. revenue threshold to Canadian dollars I use the official currency conversion for 1870, with 1 Canadian dollar being equivalent to 1.1492 U.S. dollars as reported in the online data set from Jacks and Pendakur (2010).

logarithm of labor productivity (value added divided by employment weighted by months in operation). The main coefficients of interest are associated with the indicator variables for firms' use of steam and water power. Column (1) reports estimates from a regression model including the main variables of interest and indicator variables that control for industry-specific effects. Column (2) adds controls for small firms (i.e. those having five or fewer employees), the gender composition of labor, and the population in the firm's township. Finally, column (3) includes an additional control variable for capital-intensity.

The descriptive regression results in Table 2 show a positive correlation between the use of steam and water power and labor productivity. The coefficient on the steam power variable is positive, and statistically significant at the one percent level in each specification. The coefficient on the water power variable is positive, and statistically significant at the one percent level in specifications (1) and (2), and not statistically significant in specification (3). The magnitude of water and steam power coefficients drops considerably in specification (3).[7] This is not surprising as steam and water power were embodied in physical capital. The coefficients for these variables in column (3) therefore capture the correlation between these technologies and labor productivity as distinct from other forms of capital investment.

The coefficients for the other control variables in Table 2 are consistent with the literature on the North American manufacturing sector during this era.[8] The negative and highly significant coefficient for the small firm indicator variable is similar to the results of Inwood and Keay (2012), who find evidence of returns to scale in manufacturing during this era. The positive and highly significant coefficient on the variable measuring the fraction of male employees is consistent with the findings of Goldin and Sokoloff (1982), who find that female wages were approximately 50 percent lower than male wages during the mid-nineteenth century in the U.S. manufacturing sector. Finally, the positive and highly significant coefficient for the township population variable is analogous to the results of Inwood and Keay (2005), who find evidence of a positive correlation between market size and productivity for the Canadian manufacturing sector during this era.

The final descriptive regressions in this section of the paper analyze the relationship between steam power adoption and population density. Table 3 presents the results from regressions of the natural logarithm of population density on two different measures of steam power adoption at the township level.[9] The first column shows that township population density is positively correlated with the fraction of firms using steam power, as the slope coefficient is positive and statistically significant at the one percent level. The results in the second column show that township population density is also positively correlated with the fraction of mechanized firms using steam power. The slope coefficient is positive and highly significant, and the R-squared statistic in the second specification is much higher than the first. This suggests there is an important relationship between population density and the transition from water to steam power.

The empirical analysis in this section can be summarized by three main results. First, firms using steam power were large, with value added that was nearly nine times the average for the sector. Second,

[7] The p-values at the bottom of Table 2 indicate that the difference between the water and steam power coefficients is not statistically significant in specifications (1) and (2), and statistically significant at the one percent level in specification (3).

[8] The coefficient estimates for the small firm indicator variable, the fraction of male employees, and the natural log of district population are statistically significant at the one percent level in specifications (2) and (3).

[9] Each regression is weighted by 1871 township populations.

Table 3: Descriptive Regression of Township Population Density on Steam Power Adoption

Variable	(1)	(2)
Fraction of Firms in Township using Steam	2.380***	
	(0.640)	
Fraction of Mechanized Firms in Township using Steam		2.477***
		(0.207)
Constant	-2.190***	-3.121***
	(0.115)	(0.128)
Observations	506	484
R-squared	0.027	0.229

*** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the natural logarithm of population density. Regressions are weighted by 1871 township population.

the descriptive regressions show that labor productivity was positively correlated with the use of both water and steam power. However when capital intensity is controlled for, only the coefficient on steam power remains positive and statistically significant. Third, steam power adoption is positively correlated with population density at the township level.

The descriptive analysis in this section does not identify the causal effect of steam power adoption productivity, welfare, or urbanization. In Sections 3 and 4 I develop and estimate a model of firm heterogeneity, technological adoption, and urbanization, with the aim of achieving these objectives.

3 Theoretical Model of Firm Heterogeneity, Technological Adoption, and Urbanization

This section embeds a model of firm heterogeneity and technological adoption into the new economic geography framework of Helpman (1998). The model is developed with the objective of deriving theoretical measures of the effect of mechanical steam power on welfare and urbanization, which can be estimated using the data that was described in section 2.

3.1 Model

Consider an economy with L identical households, each with one unit of labor. Households move freely between the N townships of the economy, indexed $n \in \{1, 2, ..., N\}$. In every township there is a fixed and immobile supply of residential land, H_n. There is an endogenous measure of firms in each township, M_n, that produce consumption goods that are traded freely across townships. Firms produce goods by hiring labor and choosing one of three different production technologies: non-mechanical power, water power, or steam power. These three production technologies differ in terms of their productivity and adoption cost.

Following Helpman (1998), the utility of a household in township n takes the following form:

$$u_n = q_n^{\alpha} h_n^{1-\alpha}, \quad where: \quad h_n = \frac{H_n}{L_n}, \quad q_n = \left(\int_{\omega \in \Omega} q_n(\omega)^{\rho} d\omega \right)^{\frac{1}{\rho}}, \quad \rho \equiv \frac{\sigma - 1}{\sigma}, \quad \sigma > 1, \quad 0 < \alpha < 1 \quad (1)$$

In this utility function, L_n is the endogenous population in township n, and q_n is the CES consumption index of a representative household in township n. Households are endowed with one unit of labor and own an equal share in the aggregate profits of firms residing in their township. I follow Michaels et al. (2012) in assuming that land expenditures are returned to township households as a lump sum transfer.

As in Bustos (2011), firms make a discrete choice to operate using one of the following production technologies:

$$q(\varphi, \gamma) = \varphi\gamma\ell, \quad \gamma \in \{\gamma_1, \gamma_2, \gamma_3\}, \quad \gamma_3 > \gamma_2 > \gamma_1 = 1 \tag{2}$$

The production technologies use variable labor input, ℓ, and the parameter γ differentiates the production technologies across the three different power motives. That is, firms' choice to use non-mechanical, water, or steam power is formalized by the associated choice of the production technology, $q(\varphi, 1)$, $q(\varphi, \gamma_2)$, or $q(\varphi, \gamma_3)$ respectively.[10]

Steam power's superiority over water and non-mechanical power is captured by the assumption that $\gamma_3 > \gamma_2 > 1$. It is important to empirically validate the assumption that steam increased firm level productivity relative to water power, $\gamma_3/\gamma_2 > 1$. The validity of the welfare statistic depends on this assumption, and the estimated welfare gains are increasing in the magnitude of this ratio. In section 4, I discuss the econometric framework and identification strategy that is used to derive a causal estimate of the productivity of steam relative to water power.

The production technology in equation (2) also depends on the parameter φ, which is the firm's idiosyncratic productivity type. This parameter is realized through the process summarized in Figure 1. At stage A, a prospective firm chooses to enter the industry by paying a fixed cost $w_n f_e$. All fixed costs are denominated in units of labor paid the township specific wage, w_n. After paying the entry cost, firms draw their productivity type, φ, from a Pareto CDF, $G(\varphi) = 1 - \varphi^{-\theta}$. The distribution of productivity types, $G(\varphi)$, is assumed to be *ex-ante* identical across townships. I assume that the shape parameter of the Pareto distribution, θ, satisfies $\theta > \sigma - 1$, which ensures that the ex-post average productivity of firms is finite.[11] Upon realizing their type, firms face a decision at stage C to exit or remain in the industry. At stage D, firms choose a power motive and pay a technology specific fixed cost of production. The firm's decision to use non-mechanical, water, or steam power is associated with a fixed cost of production equal to $w_n f$, $w_n \eta_{n,2} f$, or $w_n \eta_{n,3} f$ respectively. The fixed costs associated with steam and water power are modeled as varying across townships. For water power, the township specific adoption cost captures the fact that the cost of waterwheel construction depended on geographic factors. For steam power, variation in cost of adoption reflects the fact that access to steam power may have differed depending on the the location of the township. At stage E firms choose the price and output level that maximizes their profits.

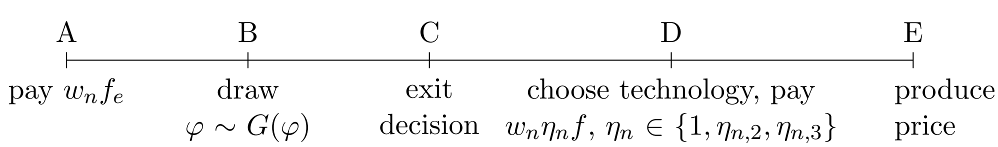


Figure 1: Firms' Timeline

[10] I normalize the productivity parameter for the non-mechanized technology to one for notational convenience, $\gamma_1 = 1$.

[11] This assumption is validated in the empirical analysis.

Depending on the model parameters, there may be townships where one or more of the three production technologies are not used in equilibrium. In 1871, there were non-mechanized firms operating in over 95 percent Ontario townships. To match this empirical regularity and simplify the presentation of the model, I assume that the cost of water and steam power are sufficiently high, such that there is a positive measure of non-mechanized firms in every township in equilibrium.

This leaves four possible cases with regards to the production technologies used in any particular township. Townships can be categorized as one of four possible types: (i) having both water and steam powered firms; (ii) having neither water nor steam powered firms; (iii) having water but not steam powered firms; (iv) having steam but not water powered firms. For the sake tractability, I assume that all townships are of types (i), (ii), or (iii).[12] In theory, it is possible to also include type (iv) townships in the theoretical framework by making appropriate assumptions on the model parameters. Unfortunately, this generalization results in welfare and urbanization statistics that can no longer be tractably estimated, given the data described in section 2.[13]

In appendix 6.1.1 I provide details on the solution to the theoretical model that has been described above. In this section I describe the unique equilibrium of the model, and explain how the theoretical framework can be used to derive statistics that measure the effect of steam power on urbanization and welfare.

The model gives rises to endogenous idiosyncratic productivity cut-offs that determine each firms' technological choice. Define $\varphi_{n,1}$, $\varphi_{n,2}$, and $\varphi_{n,3}$ as the operating, water power, and steam power cut-off productivity levels respectively. In equilibrium, the ordering of the cut-offs, $\varphi_{n,1} < \varphi_{n,2} < \varphi_{n,3}$, follows the ordering of the technological productivity parameters, $1 < \gamma_2 < \gamma_3$. Firm's with the highest productivity ($\varphi > \varphi_{n,3}$) use steam power; those with intermediate levels of productivity ($\varphi_{n,2} < \varphi < \varphi_{n,3}$) use water power; and active firms with the lowest levels of productivity ($\varphi_{n,1} < \varphi < \varphi_{n,2}$) operate without mechanized power.[14]

The equilibrium values of the water and steam power cut-offs are proportional to the operating firm cut-off:

$$\frac{\varphi_{n,2}}{\varphi_{n,1}} = \left(\frac{\eta_{n,2} - 1}{\gamma_2^{\sigma-1} - 1}\right)^{\frac{1}{\sigma-1}}, \qquad \frac{\varphi_{n,3}}{\varphi_{n,1}} = \left(\frac{\eta_{n,3} - \eta_{n,2}}{\gamma_3^{\sigma-1} - \gamma_2^{\sigma-1}}\right)^{\frac{1}{\sigma-1}}, \tag{3}$$

[12]In particular, this assumption requires the following parameter restrictions: $\eta_{n,2} - 1 > \gamma_2^{\sigma-1} - 1$, and $(\eta_{n,3} - \eta_{n,2})/(\eta_{n,2} - 1) > (\gamma_3^{\sigma-1} - \gamma_2^{\sigma-1})/(\gamma_2^{\sigma-1} - 1)$. The first assumption ensures that a positive measure of firms operate without mechanical power. The second assumption ensures that a positive measure of firms operate with water power, provided that $\eta_{n,2} < \infty$. The model is consistent with type (ii) townships, when $\eta_{n,2}$ and $\eta_{n,3}$ approach infinity. In any such township, the fixed cost of adopting either water or steam power is prohibitively expensive, resulting in a measure zero of mechanized firms. The model is consistent with townships of type (iii), when $\eta_{n,2}$ is finite, and $\eta_{n,3}$ approach infinity. In type (iii) townships, the cost of adopting steam is prohibitively expensive, which results in a no firms using steam power.

[13]In 1871, approximately 17 percent of Ontario townships in 1871 were of type (iv) (that is, having steam powered firms, but no firms using water power). In appendix 6.1.3, I derive the theoretical properties of the welfare and urbanization statistics under the more general setting that incorporates type (iv) townships. For type (iv) firms, I show that the urbanization statistic given by equation (9) will typically be a lower bound, relative to the more general framework. However, this need not necessarily be the case if there are numerous type (iv) townships. For townships of all other types, the urbanization equation (9) is an upper bound. Finally, I show that excluding type (iv) townships from the theoretical framework has an ambiguous effect on the welfare statistic, which depends on the model's parameters.

[14]Note that a township may have no firms using steam power if the fixed cost of adoption, $\eta_{n,3}$, is infinite, in which case $\varphi_{n,3}$ is also infinite. Similarly, a township may have no mechanized firms if both $\eta_{n,2}$ and $\eta_{n,3}$ are infinite, in which case $\varphi_{n,2}$ and $\varphi_{n,3}$ are infinite.

and the operating firm cut-off is given by the following expression:

$$\varphi_{n,1} = \left(\frac{f(\sigma-1)}{f_e(\theta-\sigma+1)} \left[1 + (\eta_{n,2}-1) \left(\frac{\varphi_{n,2}}{\varphi_{n,1}} \right)^{-\theta} + (\eta_{n,3}-\eta_{n,2}) \left(\frac{\varphi_{n,3}}{\varphi_{n,1}} \right)^{-\theta} \right] \right)^{\frac{1}{\theta}} \tag{4}$$

Define $\varphi_{n,1}^{'}$ as the equilibrium operating cut-off in township n when steam power is not available in the economy. Solving the model without steam power yields:

$$\varphi_{n,1}^{'} = \left(\frac{f(\sigma-1)}{f_e(\theta-\sigma+1)} \left[1 + (\eta_{n,2}-1) \left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1} \right)^{-\frac{\theta}{\sigma-1}} \right] \right)^{\frac{1}{\theta}} = \lim_{\eta_{n,3}\to\infty} \varphi_{n,1} = \lim_{\gamma_3\to\gamma_2} \varphi_{n,1} \tag{5}$$

Intuitively, when the cost of steam power adoption becomes prohibitively high, or when the productivity gain associated with steam power adoption is small relative to water power, the equilibrium operating cut-off approaches its value in the economy without steam power.

The model is closed by solving for the endogenous population, L_n, and the measure of firms, M_n, in each townships, which are given by the following expressions:

$$L_n = L \frac{H_n \varphi_{n,1}^{\frac{\sigma-1}{\sigma}\frac{\alpha}{1-\alpha}}}{\sum_{i=1}^{N} H_i \varphi_{i,1}^{\frac{\sigma-1}{\sigma}\frac{\alpha}{1-\alpha}}}, \qquad M_n = \frac{\sigma-1}{\sigma\theta fe} \frac{L_n}{\varphi_{n,1}^{\theta}} \tag{6}$$

Intuitively, a township's population is increasing in its land endowment, H_n, and the cut-off $\varphi_{n,1}$, which is the scale parameter of the township's ex-post idiosyncratic productivity distribution. More productive townships are more populated and also have higher nominal wages, as the log difference in wages between any two townships is proportional to the log difference in their operating cut-offs, $ln(w_n/w_i) = \rho ln(\varphi_{n,1}/\varphi_{i,1}),\ \forall i,n \in \{1,2,...N\}$.[15]

Finally, note from equation (4) that each township's operating cut-off is increasing the fraction of firms using steam power, $(1-G(\varphi_{n,3}))\,/\,(1-G(\varphi_{n,1})) = (\varphi_{n,3}/\varphi_{n,1})^{-\theta}$. It follows that the equilibrium township population is increasing the fraction of firms using steam power, which is consistent with the stylized facts presented in section 2. In the model, townships with higher rates of steam power adoption have higher productivity, higher nominal wages, and larger populations. However, households residing in such townships must tolerate the congestion externality that results from the fixed supply of residential land, H_n. In equilibrium, the free mobility of labor equalizes household utility between high wage densely populated townships with high rates of technological adoption, and low wage sparsely populated townships with low rates of adoption.

3.2 Welfare, Urbanization, and Technological Adoption

In this section I present measures of the effect of a new technology on welfare and urbanization that are derived from the model develop in section 3.1.[16] In section 4, I demonstrate how these measures can be estimated from the data described in section 2. These statistical measures compare welfare and

[15] See appendix 6.1.1 for the derivation of this equality.

[16] Derivations of the welfare and urbanization statistics are provided in appendix 6.1.2.

urbanization in the observed economy versus a counterfactual economy where steam power is unavailable.[17] Throughout this section I denote equilibrium values from the counterfactual economy with primes. For example, welfare in the economy without steam power is denoted $\mathbb{W}'$, and the endogenous population of township n in the counterfactual economy is denoted L_n'. From the utility function in equation (1), the logarithmic approximation to the percentage change in welfare from the introduction of steam is:

$$ln\left(\mathbb{W}'/\mathbb{W}\right) = \alpha ln\left(\frac{w_n P'}{w_n' P}\right) - (1-\alpha) ln\left(\frac{L_n}{L_n'}\right), \quad \forall n \in \{1,2,...,N\} \tag{7}$$

The free mobility of households ensures that the changes in welfare are equalized across all townships.

In this paper the theoretical and empirical measure of urbanization is the change in population density at the township level. However, since the land endowment of each township is held constant in the counterfactual exercise, the change in urbanization in each township is equivalent to the change in population. That is, the logarithmic approximation to the percentage change in urbanization in township n is: $\%\Delta\ urbanization_n \cong ln\left(L_n H_n / \left(L_n' H_n\right)\right) = ln\left(L_n / L_n'\right)$.

In appendix 6.1.2, I show that for the counterfactual exercise, the effect of steam on welfare and urbanization are summarized by the following statistics:

$$ln\left(\mathbb{W}'/\mathbb{W}\right) = \frac{\alpha}{1-\sigma} ln\left(\sum_{n=1}^{N} \frac{R_n}{R}\left(1 + \frac{R_{n,3}}{R_n}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma} - 1\right)\right)^{\frac{\sigma-1}{\theta}} \left(\frac{L_n}{L_n'}\right)^{\frac{\sigma(1-\alpha)-1}{\alpha}}\right) \tag{8}$$

$$ln\left(\frac{L_n}{L_n'}\right) = \frac{\alpha(1-\sigma)}{\theta(1-\alpha)\sigma} ln\left(1 + \frac{R_{n,3}}{R_n}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma} - 1\right)\right) + ln\left(\sum_{n=1}^{N} \frac{L_i}{L}\left(1 + \frac{R_{i,3}}{R_i}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma} - 1\right)\right)^{\frac{\alpha(\sigma-1)}{\theta(1-\alpha)\sigma}}\right) \tag{9}$$

Where R_n is the aggregate revenue in township n, and $R_{n,3}$ is the aggregate revenue of steam powered firms in township n. Given estimates of the aggregate values R_n, $R_{n,3}$ L_n and L, and parameter estimates of α, σ, γ_3/γ_2, and θ, it is possible to calculate the welfare and urbanization statistics. The empirical application in section 4 estimates these statistics using the data described in section 2 on mechanical steam power adoption in Ontario in 1871.

4 Empirical Application

4.1 Econometric Model

Estimating the urbanization and welfare statistics given by equations (8) and (9) requires estimates of the following model parameters: θ, α, the ratio γ_3/γ_2, and σ. In this section I derive two structural equations from the theoretical model that are used to estimate the first three of these four parameters. The elasticity of substitution, σ, is not identified in the econometric analysis and is parametrized based on values used in the empirical international trade and industrial organization literature.

[17] The methodology developed in this section draws from the literature on estimating the welfare gains from international trade, including Melitz and Redding (2014), Arkolakis et al. (2012), Feenstra (2010).

The Pareto shape parameter, θ, is estimated from the subset of firms using steam power across all townships. Under monopolistic competition and CES demand, the revenue of a steam powered firm in township n is: $r_n(\varphi, \gamma_3) = \sigma f w_n \gamma_3^{\sigma-1} (\varphi/\varphi_{n,1})^{\sigma-1}$.[18] Dividing this expression by the mean revenue across all steam powered firms in township n, $\tilde{r}_n(\gamma_3) = \theta\sigma f w_n \gamma^{\sigma-1} (\varphi_{n,3}/\varphi_{n,1})^{\sigma-1} / (\theta - \sigma + 1)$, normalizes the revenue function such that it is Pareto distributed and i.i.d. across townships:

$$F(r) = \begin{cases} 0 & \text{if } \varphi < \varphi_{n,3} \\ 1 - \left(\frac{\theta}{\theta-\sigma+1}\right)^{\frac{-\theta}{\sigma-1}} (r)^{\frac{-\theta}{\sigma-1}} & \text{if } \varphi \geq \varphi_{n,3} \end{cases}, \qquad r \equiv \frac{r_n(\varphi, \gamma_3)}{\tilde{r}_n(\gamma_3)} = \frac{\theta - \sigma + 1}{\theta} \left(\frac{\varphi}{\varphi_{n,3}}\right)^{\sigma-1} \tag{10}$$

That is, the random variable r is Pareto distributed with scale parameter $(\theta-\sigma+1)/\theta$, and shape parameter $\theta/(\sigma-1)$.

For the subset of firms using steam power, the CDF $F(r)$ can be re-written:

$$ln(r) = ln\left(1 - \frac{\sigma - 1}{\theta}\right) - \frac{\sigma - 1}{\theta} ln(1 - F(r)), \tag{11}$$

The empirical specification corresponding to equation (11) is:

$$ln(r_i) = \zeta_0 + \zeta_1 ln(1 - \hat{F}_{i,n}) + \epsilon_{i,n}, \qquad r_i \equiv \frac{r_{i,n}(\varphi_{i,n}, \gamma_3)}{\tilde{r}_n(\gamma_3)} \tag{12}$$

I follow Head et al. (2014) in estimating equation (12) by the QQ regression methodology introduced by Kratz and Resnick (1996). The variable $r_{i,n}(\varphi_{i,n}, \gamma_3)$ is specified as the value added of steam powered firm i, in township n; $\tilde{r}_n(\gamma_3)$ is the average value added of steam powered firms in township n; and $\hat{F}_i$ is the empirical CDF for the variable r_i. Regression equation (12) is estimated using the sub sample of firms using steam power.

Given an estimate of the elasticity of substitution, the shape parameter, θ, is identified by the coefficient $\zeta_1 = -(\sigma - 1)/\theta$. Hsieh and Klenow (2009) note that estimates of σ in the empirical trade and industrial organization literature typically range from 3 to 10 for the manufacturing sector. For my analysis, I follow Melitz and Redding (2013) and Head et al. (2014) in using a benchmark parametrization of $\sigma = 4$ and explore the robustness of this assumption in the sensitivity analysis.

I now derive an econometric framework for estimating the parameters α and γ_3/γ_2 from the theoretical model. The revenue of any firm in township n can be written: $r_n(\varphi, \gamma) = \sigma f w_n \gamma^{\sigma-1} (\varphi/\varphi_{n,1})^{\sigma-1}$. Next, from the utility function in equation (1), the the nominal township wage can be written: $w_n = (P^\alpha \mathbb{W})^{1/\alpha} (L_n/H_n)^{(1-\alpha)/\alpha}$. Combining these two expressions yields:

$$r_n(\varphi, \gamma) = \sigma f (P^\alpha \mathbb{W})^{1/\alpha} (L_n/H_n)^{(1-\alpha)/\alpha} \gamma^{\sigma-1} (\varphi/\varphi_{n,1})^{\sigma-1} \tag{13}$$

From the equilibrium solution to the theoretical model, the measure of firms using mechanized power

[18] Here I have made use of the result that a firm with idiosyncratic productivity equal to $\varphi_{n,1}$ makes zero profits, which implies: $r_n(\varphi_{n,1}, 1) = \sigma f w_n$.

is: $(M_{n,2}+M_{n,3})/M_n = (1-G(\varphi_{n,2}))\,/\,(1-G(\varphi_{n,1})) = (\varphi_{n,2}/\varphi_{n,1})^{-\theta}$.[19] Combining this expressions with equation (13), and considering only the subset of firms using mechanical power yields:

$$ln(r_n(\varphi,\gamma)) = ln\left(\sigma f\left(P^{\alpha}\mathbb{W}\right)^{1/\alpha}\gamma_2^{\sigma-1}\right) + \frac{1-\alpha}{\alpha}ln\left(\frac{L_n}{H_n}\right) + 1_{\gamma_3}(\sigma-1)ln\left(\frac{\gamma_3}{\gamma_2}\right)$$
$$-\frac{\sigma-1}{\theta}ln\left(\frac{M_{n,2}+M_{n,3}}{M_n}\right) + (\sigma-1)ln\left(\frac{\varphi}{\varphi_{n,2}}\right), \qquad (14)$$

where 1_{γ_3} is an indicator function that takes a value of one if the firm uses steam and zero otherwise. The empirical specification corresponding to equation (14) is:

$$ln(r_{i,n}(\varphi_{i,n},\gamma)) = \beta_0 + \beta_1 ln\left(\frac{L_n}{H_n}\right) + \beta_2 1(steam_{i,n}) + \beta_3 ln\left(\frac{M_{n,mech}}{M_n}\right) + u_{i,n}, \qquad (15)$$

When presenting the results in section 4.3, I refer to equation (15) as the value added regression, as the dependent variable, $r_{i,n}(\varphi_{i,n},\gamma)$, is specified as value added. The variable L_n/H_n is the population density of township n; $1(steam_{i,n})$ is a steam power indicator variable that takes a value of one if firm i uses steam and zero otherwise; and $(M_{n,mech})/M_n$ is the fraction of mechanized firms in township n. The error term is defined: $u_{i,n} = (\sigma-1)ln(\varphi_{i,n}/\varphi_{n,2})$. Implicitly dividing $\varphi_{i,n}$ by $\varphi_{n,2}$ provides the necessary normalization so that the error term is i.i.d. across townships. The value added regression in equation (15) is estimated with the sub sample of firms using mechanized power (i.e. water or steam power).

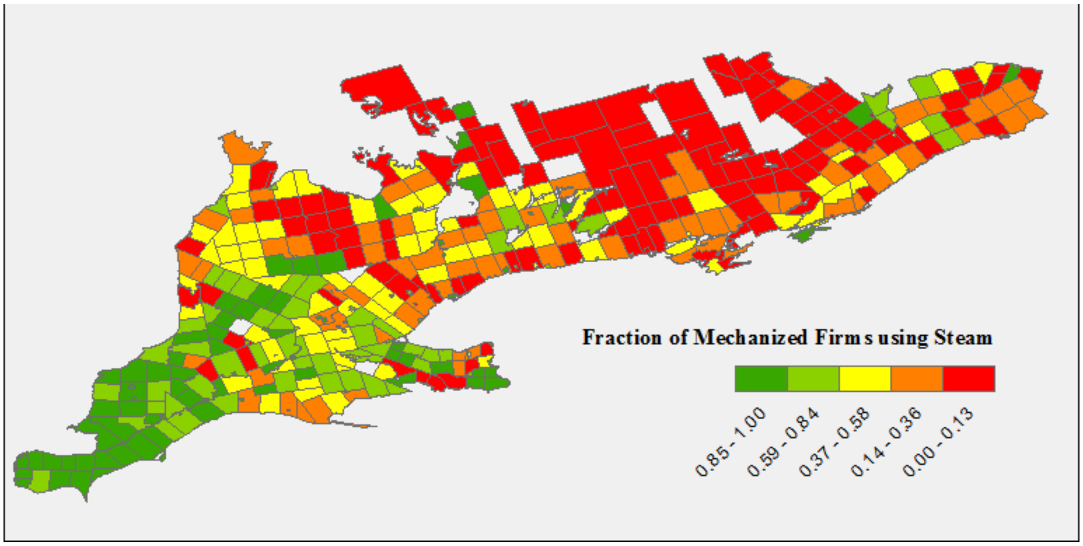
Estimates of the model parameters α and the ratio γ_3/γ_2 can be derived from estimate of the regression coefficients β_1 and β_2 respectively.[20] Section 4.3 also present results for a specification of equation (15) where the natural logarithm of township population, L_n, and area, H_n, enter separately. I report tests of the parameter restriction that the coefficients on $ln(L_n)$ and $ln(H_n)$ should be equal in absolute value and opposite in sign, as is implied by equation (15).

4.2 Identification

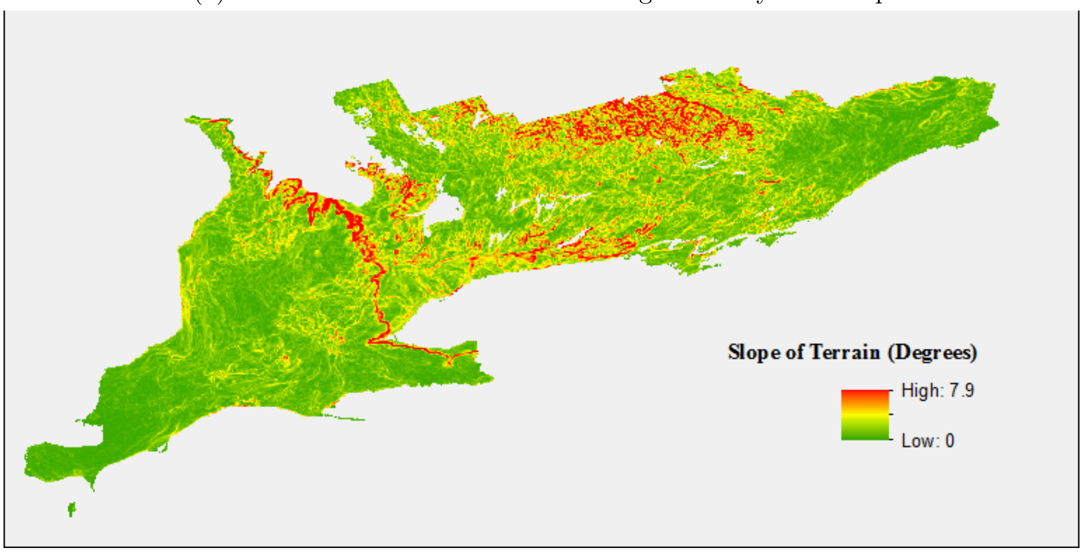
If a firm's productivity type is positively correlated with the adoption decision, then the coefficient β_2 in equation (15) will be biased upwards if estimated by OLS regression. A consistent estimate of β_2 can be obtained using instrumental variables (IV) estimation. The identification strategy used in this paper exploits exogenous variation in the average terrain slope within a township. The idea for this identification strategy comes from Bishop and Muñoz-Salinas (2013), who study the location of historic watermills in England and Scotland. Bishop and Muñoz-Salinas (2013) find that terrain steepness is an important factor in explaining the historic location of watermills. The authors explain this result by noting that upstream dams were required to create a reservoir to power watermills. In steeper channels of a river, the height

[19]See appendix 6.1.1 for details on the equilibrium solution to the model. Note that this equation is applicable if water power is used in the township, regardless of whether or not steam power is used. If water power is *not used*, then the fraction of mechanized firms is: $M_{n,3}/M_n = (1-G(\varphi_{n,3}))\,/\,(1-G(\varphi_{n,1})) = (\varphi_{n,3}/\varphi_{n,1})^{-\theta}$. In this case, the regression equation given by (15) is still correctly specified. The inclusion of the fraction of mechanized firms in regression equation (15) implicitly normalizes the error term so that it is i.i.d. across townships. To see this, note that in townships without water power, $(M_{n,2}+M_{n,3})/M_n = M_{n,3}/M_n = (\varphi_{n,3}/\varphi_{n,1})^{-\theta}$. This implies that the error term is $u_{i,n} = (\sigma-1)ln(\varphi_{i,n}/\varphi_{n,3})$. Dividing $\varphi_{i,n}$ by $\varphi_{n,3}$ provides the necessary normalization so that the error term is i.i.d. across townships.

[20]Identifying the ratio γ_3/γ_2 requires an estimate of the elasticity of substitution, σ.



(a) Fraction of Mechanized Firms using Steam by Township



(b) Terrain Slope, Calculated from Digital Elevation Model

Figure 2: Ontario Townships in 1871 and Terrain Slope

and cost of building upstream dams was lower. Bishop and Muñoz-Salinas (2013) argue that this is why watermills were commonly found in areas with steeper terrain.

The negative correlation between terrain slope and steam power adoption has also been noted by economic historians such as Peter Temin. Temin (1966) notes that starting in 1822, Louisiana sugar mills were among the earliest adopters of mechanical steam power. Despite the fact that the Mississippi river and many of its tributaries pass through the state, Temin notes that "the land was too flat to provide suitable waterpower sites" (p. 202).

In the context of the theoretical model, terrain slope lowers the the cost of water power adoption, $\eta_{n,2}$, thereby *decreasing* the probability of steam power adoption in townships with steeper terrain. Importantly, the fixed cost of water power does not enter in the theoretical regression equation, (14). Therefore, the instrument is orthogonal to the error term and thus satisfies the standard IV exclusion restriction in theory.

To generate the instrument I calculate the terrain slope using a digital elevation model (DEM) and data from the NASA Shuttle Radar Topography Mission (SRTM).[21] I then define the instrument as the natural log of the average slope within each township.

Panel (a) of Figure 2 presents a map of the 1871 Ontario townships.[22] This map uses graduated symbology to depict the fraction of mechanized firms using steam power in each township.[23] Panel (b) shows the terrain slope for Ontario, as calculated by the digital elevation model. The negative correlation between terrain slope and the use of steam power can be seen visually Figure 2, and provides evidence in support of the identification strategy. This is consistent with the large first stage F-statistics that are reported in the IV results in section 4.3.

4.3 Empirical Results

The results from estimating the value added regression, (15), and the QQ-regression, (12), are presented in Table 4. I estimate the value added regression by OLS and instrumental variables estimation. The IV estimator uses the *Procedure 18.1* approach proposed by Wooldridge (2002), which leverages the binary nature of the endogenous steam power indicator variable.[24] Standard errors in each value added regression are clustered at the township level.

Two specifications of regression equation (15) are estimated. The first model allows the natural logarithm of township population and area to enter the regression equation separately. The second model includes the natural logarithm of population density, thus constraining the coefficients on the population and area variables to be equal in magnitude and opposite in sign. The p-value from the test of this parameter restriction is reported at the bottom of the value added regression results in Table 4. These p-values are large enough that the parameter restriction cannot be rejected at any reasonable level of significance. Given that this parameter restriction is also implied by the theoretical model, I select the restricted model as the preferred specification.

Estimates of α are derived from the coefficient estimates on $ln(Population_n)$ and $ln(Population\ Density_n)$, in the first and second specification respectively. Recall from the theoretical model that $1-\alpha$ is the share of household income expenditure on residential land. Using contemporary data for the U.S., Davis and Ortalo-Magne (2011) find that residential expenditures account for approximately 25 percent of total household expenditures, which implies a value of $\alpha = 0.75$. This is very close to the estimated values of α reported in Table 4.

The coefficient estimates for the steam power indicator variable are positive and statistically significant

[21]One kilometer square SRTM digital elevation model data was downloaded from the Consultative Group for International Agricultural Research - Consortium for Spatial Information (CGIAR-CSI) website: http://www.cgiar-csi.org/

[22]Historic township boundaries are reconstructed from the CANIN71 GIS boundary files for the 1871 census sub-districts (CSDs). In the case of cities, the 1871 Census enumerated neighborhood wards as individual CSDs. Using GIS analysis, I dissolve ward boundaries to re-construct the geographic boundaries of cities. For all other CSDs, township boundaries coincide with CSD boundaries. I thank Josh MacFadyen for providing the geo-spatial files for the 1871 CSDs.

[23]Non-mechanized township (i.e. those without firms using steam or water) are not shown in Panel (a). However, these townships are incorporated in the welfare and urbanization analysis in section 4.3.

[24]The first step in Wooldridge's *Procedure 18.1* IV estimator is to run a probit regression of the steam power indicator variable on the full set of instruments and the control variables from the equation (15). The fitted probabilities from the probit regression are then used as the instrument in estimating equation (15) by two-stage least squares. Importantly, Wooldridge's *Procedure 18.1* estimator remains consistent, and the standard errors of the coefficient estimates are asymptotically valid, even when probit regression is miss-specified.

at the one percent level in each of the value added regression specifications. These coefficient estimates and the parametrized value of the elasticity of substitution, $\sigma = 4$, are used to calculate estimates of the productivity gain from steam relative to water, $(\gamma_3 - \gamma_2)/\gamma_2$. The magnitude of the productivity estimates decline between the OLS and IV estimators. This is consistent with the idea that positive selection results in an upward bias in the OLS estimate of the effect of steam power on productivity. The causal interpretation of the IV estimates are that the gain from switching from water to steam power was a 22.0, and 21.9 percent percent increase in firm level productivity under the first and second specifications respectively.[25] While the magnitude of these estimates are large, they are not unreasonable when they are interpreted in light of the fact that the average value added of firms using steam power was more than double that of water powered firms.

The F-statistics reported at the bottom of the value added regressions provide a test of the strength of the instrument. For the *Procedure 18.1* IV estimator, Wooldridge (2002) advocates reporting the standard IV weak instruments test. In particular, the F-statistics in Table 4 are calculated from the output of a linear regression of the steam power variable on the exogenous explanatory variables and the instrument, $ln(slope_n)$. The F-statistic tests the null hypothesis that the coefficient on the instrument is equal to zero. The F-statistic is sufficiently large to soundly reject the null hypothesis of weak-identification at any reasonable level of significance. A full set of results from this auxiliary regression are reported in Table 5, in appendix 6.2.1. The fact that the coefficient on the slope variable is the expected sign (negative), and highly significant, provides further evidence in support of the identification strategy. The strong positive correlation between steam power adoption and population density is also consistent with the theoretical framework, and the descriptive analysis presented in section 2.

The QQ regression results are reported in the middle panel of Table 4.[26] The QQ regression slope coefficient is less than one, which is consistent with theoretical modeling assumption that $\theta > \sigma - 1$. The estimate of the Pareto shape parameter, θ, uses the parametrized value of $\sigma = 4$, and the slope coefficient estimate from the QQ regression. The estimated value the shape parameter is slightly lower than values used in the international trade literature, indicating a slightly greater degree of firm heterogeneity.[27]

The bottom panel of Table 4 reports the welfare estimates. The estimates measure the welfare gains in moving from a counterfactual economy without steam power to the observed economy. Standard errors for the welfare statistic are calculated using a non-parametric bootstrap, clustered at the township level.[28]

Under the preferred specification, IV 2, the results indicate that the introduction of steam power resulted in a 5.83 percent increase in welfare. As previously noted, the framework developed in this paper

---

[25]These estimates are statistically significant at the one percent level.

[26]The QQ regression estimates do not depend on the results from the value added regression, and are therefore only listed in the first column of the table. Standard errors for the QQ regression are clustered at the township level.

[27]For example, Melitz and Redding (2013) calibrate a model of heterogeneous firms in trade using values of $\theta = 4.25$ and $\sigma = 4$.

[28]I follow Cameron and Miller (2015) in estimating the cluster-bootstrap variance matrix for the coefficients of the value added and QQ regressions. In particular, I estimate the value added and QQ regressions using 1000 bootstrap replications. Each of the 1000 bootstrap samples is constructed by re-sampling (with replacement) from the original township clusters. The welfare statistic is written as a function of the observed data and regression coefficients from the value added and QQ regressions. The standard error for the welfare statistic is then calculated using the delta method, and the cluster-bootstrap variance matrix for the coefficients of the value added and QQ regressions.

Table 4: Estimation Results

Value Added Regression	OLS 1	OLS 2	IV 1	IV 2
$ln(Population_n)$	0.342***		0.336***	
	(0.0325)		(0.0362)	
$ln(Land\ Area_n)$	-0.294***		-0.303***	
	(0.0195)		(0.0215)	
$ln(Population\ Density_n)$		0.301***		0.308***
		(0.0185)		(0.0217)
Steam Power	0.651***	0.654***	0.598***	0.595***
	(0.0565)	(0.0564)	(0.194)	(0.204)
$ln(M_{n,mech}/M_n)$	0.0637	0.0617	0.0613	0.0599
	(0.0600)	(0.0600)	(0.0610)	(0.0614)
Model Parameter Estimates				
$(\gamma_3 - \gamma_2)/\gamma_2$	0.242***	0.244***	0.220***	0.219***
	(0.0234)	(0.0234)	(0.0790)	(0.0830)
α	0.745***	0.769***	0.749***	0.764***
	(0.0180)	(0.0109)	(0.0203)	(0.0127)
SIC Industry Controls	Yes	Yes	Yes	Yes
Observations	4,165	4,165	4,097	4,097
R-squared	0.308	0.308	0.194	0.194
p-value: $ln(Population_n)$				
$= -ln(Land\ Area_n)$	0.150		0.337	
F-Statistic			77.80	64.63
QQ Regression				
$ln(1 - \hat{F}_{i,n})$	-0.953***			
	(0.0161)			
Model Parameter Estimates				
θ	3.148***			
	(0.0533)			
Observations	2,028			
R-squared	0.662			
Welfare Estimates				
$ln(\mathbb{W}/\mathbb{W}')$	0.0600***	0.0636***	0.0564***	0.0583***
	(0.00594)	(0.00592)	(0.0180)	(0.0191)

Standard errors are clustered at the township level. *** p<0.01, ** p<0.05, * p<0.1. Coefficient estimates for industry indicator variables are not reported but are available upon request. The estimates of $(\gamma_3 - \gamma_2)/\gamma_2$, θ, and $ln(\mathbb{W}/\mathbb{W}')$ are calculated using a parametrized value of $\sigma = 4$. The p-value at the bottom of the value added regression results tests the hypothesis that the coefficient for the variable $ln(Land\ Area_n)$ is equal in absolute value, and opposite in sign, to the coefficient for $ln(Population_n)$. Standard errors for α, $(\gamma_3 - \gamma_2)/\gamma_2$, and θ are calculated using the delta method. Standard errors for $ln(\mathbb{W}/\mathbb{W}')$ are calculated by the non-parametric bootstrap.

draws on the literature on heterogeneous firms in international trade. The economic significance of the estimated welfare gains from steam power adoption can be appreciated by comparing the results in Table 4 with contemporary estimates of the welfare gains from trade. For example, Melitz and Redding (2013) calibrate a model of heterogeneous firms in trade using moments from the U.S. data and estimate the welfare gains from opening a closed economy to trade. Their baseline calibration suggests that the welfare gains from trade are between 2 and 3 percent. Thus, the estimated welfare gains from steam power adoption are roughly twice as large as contemporary estimates of the welfare gains from international trade.

The estimates of the productivity and welfare gains from steam power adoption are sensitive to the parametrized value of the elasticity of substitution, σ. In appendix 6.2.2, Figures 4 and 5 explore the robustness of the baseline results by plotting the estimated productivity and welfare gains when alternative values of σ ranging from 3 to 8 are used. The estimated productivity gains are decreasing in σ, with the point estimates ranging from 34.7 percent to 8.9 percent. The estimated welfare gains from steam power adoption are also decreasing in σ and ranges from a high of 9.4 percent to a low of 2.2 percent.

4.4 Evaluating the Contribution of Steam to Urbanization

The final empirical exercise of this paper evaluates the contribution of mechanical steam power to urbanization. I compare the model predicted changes in township populations from the introduction of steam, to the observed population changes for Ontario townships over the period 1861-1871.[29] Population data at the Census sub-district (CSD) level are sourced from the 1860-1861 and 1870-1871 Censuses of Canada. I combine the digitized boundary files for the 1870-1871 CSDs with historical maps from the era to create a population data set for 440 townships, that had constant geographic boundaries from 1861-1871.[30] Township population growth is therefore approximately equal to township population density growth, notwithstanding any minor CSD boundary changes that are not accounted for in my analysis.

The model predicted changes in township population is calculated from equation (9), using the preferred parameter estimates from Table 4, and the parametrized value of $\sigma = 4$. Modeled township population change is equal to population density change, since the land endowment of each township, H_n, is held constant in the counterfactual economy without steam.[31]

Figure 3 plots the model predicted and observed growth in population for Ontario townships during the period 1861-1871. The figure also includes the fitted line and estimation results from an OLS regression of

[29] It would be ideal to use township population growth over a longer period, since steam power was introduced in the Canadian economy earlier than 1861. Unfortunately, changes in the Ontario Census sub-district boundaries prior to 1861 make this infeasible. The 1861 Census of Canada reports motive of power (water or steam) for mechanized mills in a number of industries, including flour, grist, oatmeal, and saw mills. These industries were some of the earliest adopters of steam power in the Canadian manufacturing sector. The adoption rate of mechanical steam power in these industries jumped from 25 percent in 1861 to 37 percent in 1871. The rapid pace of diffusion during this era motivates the comparison of township population growth from this period, to the model predicted growth resulting from the introduction of steam power.

[30] This involved re-amalgamating CSDs that were enumerated jointly in 1861, and subsequently enumerated individually in 1871, or vice versa. A number of CSDs are excluded from the data set, when it is not possible to construct a reliable concordance between 1861 and 1871. In particular, northern townships whose boundaries were expanding with land exploration at the frontier are excluded. The dataset is completed by aggregating the two Censuses CSD population values to the township level, for each of the 440 townships with constant geography. I thank Marvin McInnis for providing township level maps and advice on the concordance of Ontario townships between 1861 and 1871.

[31] The modeled township population values are aggregated or dropped in manner that is consistent with the observed constant geography township data set. This reduces the number of modeled township observations from 506 to 440.

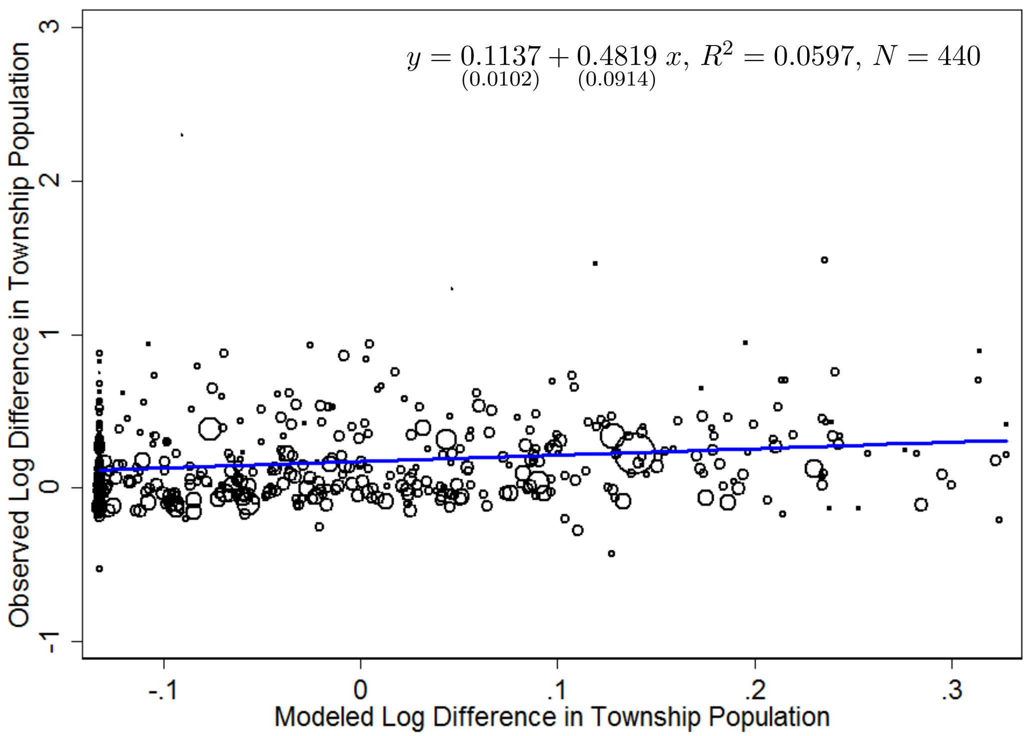


Figure 3: Township Population Growth in Ontario, 1861-1871

Regression is weighted by 1861 township population. Townships are re-constructed to have a constant geography over the period 1861-1871. By implication, $\%\Delta$ *Population* $\cong$ $\%\Delta$ *Population Density*.

observed growth on modeled growth.[32] As noted above, for both series population change is approximately equal to population density change, which is the measure of urbanization used in the paper.

The modeled change in township urbanization is qualitatively consistent the observed data, as the slope coefficient is positive and statistically significant at the one percent level. However, the R-squared statistic for the regression indicates that the model predicted urbanization only explains 5.97 percent of the observed variation. Therefore, the results in Figure 3 suggest that the introduction of steam power had a positive but economically small effect on urbanization during the mid-nineteenth century in Ontario.

The modeled effect of steam on urbanization occurs through *population re-sorting*, in a static comparison between the observed and counterfactual economies. During the period 1861-71 there were numerous other factors affecting population density growth in Ontario townships, such as structural adjustment from agriculture to manufacturing, immigration, and railroad development. The introduction of steam power may have also affected urbanization through its interaction with each of these phenomenon. In this respect, the results in Figure 3 should be interpreted as measuring the direct effect of the introduction of steam on urbanization, through population re-sorting.

5 Conclusion

This paper develops a theoretical framework for measuring the effect of a new technology on welfare and urbanization. The model is used to derive two statistics that summarize the effect of technological adoption on welfare, as measured by household real income, and urbanization, as measured by population density. I derive an econometric framework from the theoretical model, and illustrate how the welfare and urbanization statistics can be estimated from the econometric parameter estimates. The theoretical and empirical framework provide a tractable methodology for studying the interaction between economic geography, firm heterogeneity and technological adoption in a general equilibrium setting.

In addition to the theoretical contributions of the model, this paper addresses an open question in the literature regarding the effect of steam power on welfare and urbanization in the nineteenth century. The results indicate that the introduction of steam increased welfare by 5.8 percent. The magnitude of this estimate can be contextualized by drawing comparison to the international trade literature, which uses a similar methodology to measure the gains from trade. By this comparison, the estimated welfare gains from steam power are large, as Melitz and Redding (2013) estimate that the welfare gains from trade are between 2 and 3 percent. The research of Crafts (2004) on steam power diffusion in Britain during the nineteenth century provides another benchmark for evaluating the magnitude of the welfare gains that are estimated in this paper. Crafts (2004) finds that that the contribution of steam to growth in Britain peaked over the period 1850-1870 at only 0.41 percent per year. Acknowledging the significant differences in methodology, a qualitative comparison suggests that the welfare gains from steam estimated in this paper are much larger than what is implied by the results of Crafts (2004).

The large magnitude of the welfare gains may be partially explained by the fact that the theoretical and empirical framework assumes that manufacturing is the only sector in the economy. Urquhart (1986) estimates that the manufacturing sector represented 21.7 percent of Canadian GDP in 1871. If the esti-

[32]The regression and scatter plot points are weighted by 1861 township population.

mated productivity effect of steam power in manufacturing is representative of the overall effect for the aggregate economy, then the estimated welfare effects in this paper are representative of the aggregate gains. However, to the extent that there were sectors (notably agriculture) that may have benefited less from steam power, the estimated welfare gains in this paper provide and upper bound to the aggregate gains.

The estimated effect of steam on urbanization is lower than previous estimates in the literature on this topic. For example, Kim (2005) estimates that the diffusion of steam in the mid to late nineteenth century in the U.S. resulted in no more than an 8-10 percent increase in urbanization. In this paper, I find that the model predicted effect of steam power can explain approximately 6.0 percent of the variation in township population density growth in Ontario, during the period 1861-71. The model predicted effect should be interpreted as the direct effect of steam power on urbanization through population resorting, acknowledging that the introduction of steam may have interacted with other dynamic processes that contributed to urbanization. For example, the diffusion of steam power may have interacted with structural adjustment from agriculture to manufacturing, the expansion of the rail network, and immigration.

In particular, the relationship between steam power diffusion, structural adjustment, and urbanization, is an important topic for future research. Michaels et al. (2012) use a similar economic geography framework to the model developed in this paper, and find that structural adjustment from agricultural to manufacturing affected urbanization in the U.S. during the period 1880-2000. It is possible that the introduction of steam power may have accelerated the pace of structural adjustment, if steam power provided greater productivity gains in manufacturing relative to agriculture. Expanding the framework of this paper to include multiple sectors would further understanding of the potential effect of steam power on urbanization through this indirect channel. In addition, extending the theoretical framework to incorporate multiple sectors would improve the precision of the measured welfare gains, by allowing for heterogeneity in the effect of steam power on productivity across different sectors of the economy.

References

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare (2012) 'New trade models, same old gains?' *American Economic Review* 102(1), 94–130

Atack, Jeremy (1979) 'Fact or fiction? the relative costs of steam and water power: a simulation approach.' *Explorations in Economic History* 16(4), 409437

Atack, Jeremy, and Fred Bateman (1999) 'Nineteenth-century U.S. industrial development through the eyes of the Census of Manufactures.' *Historical Methods* 32(4), 177–188

Atack, Jeremy, Fred Bateman, and Thomas Weiss (1980) 'The regional diffusion and adoption of the steam engine in American manufacturing.' *Journal of Economic History* 40(2), 281–208

Atack, Jeremy, Fred Batemen, and Robert A. Margo (2008) 'Steam power, establishment size, and labor productivity growth in nineteenth century American manufacturing.' *Explorations in Economic History* 45(2), 185–198

Bishop, Paul, and Esperanza Muñoz-Salinas (2013) 'Tectonics, geomorphology and water mill location in scotland, and the potential impacts of mill dam failure.' *Applied Geography* 42, 195–205

Bloomfield, Gerald, and Elizabeth Bloomfield (1989) 'Waterwheels and steam engine in Ontario: industrial power reported in the 1871 manuscript census.' *Scientia Canadensis: Canadian Journal of the History of Science, Technology and Medicine* 13(1), 3–38

Bustos, Paula (2011) 'Trade liberalization, exports, and technology upgrading: evidence on the impact of MERCOSUR on Argentinian firms.' *American Economic Review* 101(1), 304–340

Cameron, A. Colin, and Douglas L. Miller (2015) 'A practiioner's guide to cluster-robust inference.' *Jounral of Human Resource* 50(2), 317–372

CANIND71 (2008) 'Canadian industry in 1871 project.' *University of Guelph, Ontario, http://www.canind71.uoguelph.ca/*

Crafts, Nicholas (2004) 'Steam as a general purpose technology: a growth accounting perspective.' *The Economic Journal* 114(495), 338351

Davis, Morris A., and François Ortalo-Magne (2011) 'Household expenditures, wages, rents.' *Review of Economic Dynamics* 14(2), 248–261

Feenstra, Robert C. (2010) 'Measuring the gains from trade under monopolistic competition.' *Canadian Journal of Economics* 43(1), 1–28

Goldin, Claudia, and Kenneth Sokoloff (1982) 'Women, children, and industrialization in the early republic: evidence from the manufacturing censuses.' *Journal of Economic History* 42(2), 741–774

Head, Keith, Thierry Mayer, and Mathias Thoenig (2014) 'Welfare and trade without Pareto.' *American Economic Review: Papers & Proceedings* 104(5), 310–316

Helpman, Elhanan (1998) 'This size of regions.' In *Topics in Public Economics: Theoretical and Applied Analysis,* ed. David Pines, Efraim Sadka, and Itzhak Zilcha (Cambridge University Press)

Hsieh, Chang-Tai, and Peter J. Klenow (2009) 'Misallocation and manufacturing TFP in China and India.' *The Quarterly Journal of Economics* CXXIV(4), 1403–1448

Hunter, Louis C. (1985) *A History of Industrial Power in the United States, 1780-1930. Volume Two: Steam Power.* (Greenville, DE: University Press of Virginia)

Inwood, Kris (1995) 'The representation of industry in the Canadian Census, 1871-1891.' *Histoire Sociale / Social History* 28(56), 347–374

Inwood, Kris, and Ian Keay (2005) 'Bigger establishments in thicker markets: can we explain early productivity differentials between Canada and the United States.' *Canadian Journal of Economics* 38(4), 1327–1363

— (2008) ‘The devil is in the details: assessing early industrial performance across international borders using late nineteenth and century North American manufacturers as a case study.’ *Cliometrica* 2(2), 85–117

— (2012) ‘Diverse paths to industrial development: Evidence from late-nineteenth-century Canada.’ *European Review of Economic History* 16(3), 311–333

Jacks, David S., and Krishna Pendakur (2010) ‘Global trade and the maritime transport revolution.’ *Review of Economics and Statistics* 92(4), 745–755

Jovanovic, Boyan, and Peter L. Rousseau (2005) ‘General purpose technologies.’ In *Handbook of Economic Growth,* ed. Philippe Aghion and Steven N. Durlauf (Elsevier) chapter 18, pp. 1181–1224

Kim, Sukkoo (2005) ‘Industrialization and urbanization: Did the steam engine contribute to the growth of cities in the united states?’ *Explorations in Economic History* 42, 586–598

Kratz, Marie, and Sidney I. Resnick (1996) ‘The QQ-estimator and heavy tails.’ *Stochastic Models* 12(4), 699–724

Krugman, Paul (1995a) *Geography and trade* (Cambridge: MIT Press)

— (1995b) ‘Increasing returns and economic geography.’ *Journal of Political Economy* 99(3), 483–499

Lileeva, Alla, and Daniel Trefler (2010) ‘Improved access to foreing markets raises plant-level productivity ... for some plants.’ *Quarterly Journal of Economics* 125(3), 1051–1099

Melitz, Marc J., and Stephen J. Redding (2013) ‘Firm heterogeneity and aggregate welfare.’ *National Bureau of Economic Research Working Paper* 18919, 1–34

— (2014) ‘Heterogeneous firms and trade.’ *Handbook of International Economics* 4, 1–54

Michaels, Guy, Frerdinand Rauch, and Stephen J. Redding (2012) ‘Urbanization and structural transformation.’ *The Quarterly Journal of Economics* 127(2), 535–586

Redding, Stephen J. (2012) ‘Goods trade, factor mobility and welfare.’ *NBER Working Paper*

Rosenberg, Nathan, and Manuel Trajtenberg (2004) ‘A general-purpose technology at work: The corliss steam engine in the late- nineteenth-century United States.’ *Journal of Economic History* 64(1), 61–99

Temin, Peter (1966) ‘Steam and waterpower in the early nineteenth century.’ *The Journal of Economic History* 26(2), 187–205

Urquhart, M. C. (1986) ‘New estimates of gross national product, Canada, 1870-1926: Some implications for Canadian development.’ In *Long-Term Factors in American Economic Growth,* ed. Stanley L. Engerman and Robert E. Gallman (University of Chicago Press) pp. 9–94

Wooldridge, Jeffrey M. (2002) *Econometric Analysis of Cross Section and Panel Data.* (Massachusetts Institute of Technology)

Yeaple, Stephen R. (2005) 'A simple model of firm heterogenity, international trade, and wages.' *Journal of International Economics* 65(1), 1–20

6 Appendices

6.1 Theoretical Appendix

6.1.1 Solving the Model

This appendix provides details on the equilibrium solution to the model presented in section 3.1.

Utility maximization requires that households spend a fraction, $1-\alpha$, of their income on residential land. Define ν_n as the total income of a representative household in township n. It follows that the total household income in township n can be written:

$$\nu_n L_n = w_n L_n + (1-\alpha)\nu_n L_n \Longrightarrow \alpha \nu_n L_n = w_n L_n$$

The equation on the right shows that in each township, the total expenditure on goods of all varieties, $\alpha\nu_n L_n$, equals the total wage income. Therefore, the aggregate CES demand for variety ω, with price $p(\omega)$ is:

$$q(\omega) = RP^{\sigma-1}p(\omega)^{-\sigma}, \qquad P \equiv \left(\int_{\omega\in\Omega} p(\omega)^{1-\sigma} d\omega\right)^{\frac{1}{1-\sigma}},$$

where $R = \sum_{n=1}^{N} w_n L_n$ is aggregate labor income.

I now solve the firm's problem by working backward from stage E. Monopolistic competition and CES demand implies that the price chosen by a profit maximizing firm of type φ is: $p_n(\varphi,\gamma) = w_n/\left(\rho\gamma\varphi\right)$. Combining the firm's pricing rule with household demand yields the following revenue and profit functions:

$$r_n(\varphi,\gamma) = RP^{\sigma-1}p_n(\varphi,\gamma)^{1-\sigma}, \qquad \pi_n(\varphi,\gamma) = \frac{r_n(\varphi,\gamma)}{\sigma} - w_n\eta_n f \tag{16}$$

At stage D, firms choose a power motive that affects profitability through the productivity parameter, γ, and through the fixed cost, $\eta \in \{1, \eta_{n,2}, \eta_{n,3}\}$. If steam power is used in a township, then there is a unique idiosyncratic productivity cut-off, $\varphi_{n,3}$, such that a firm with this productivity level is just indifferent with respect to using steam or water power.[33] That is, $\pi_n(\varphi_{n,3},\gamma_3) = \pi_n(\varphi_{n,3},\gamma_2)$, which from the profit function in equation (16) implies:

$$\varphi_{n,3} = \frac{w_n}{\rho P}\left(\frac{\sigma f w_n\left(\eta_{n,3}-\eta_{n,2}\right)}{R\left(\gamma_3^{\sigma-1}-\gamma_2^{\sigma-1}\right)}\right)^{\frac{1}{\sigma-1}}$$

If water power is used in a township, then there is a unique idiosyncratic productivity cut-off, $\varphi_{n,2}$, such that a firm with this productivity level is just indifferent with respect to using water power and operating without mechanized power.[34] That is, $\pi(\varphi_{n,2},\gamma_2) = \pi(\varphi_{n,2},1)$, which from the profit function in equation

[33] Townships with no steam powered firms are captured in the model by allowing $\eta_{n,3}$ to approach infinity, in which case $\varphi_{n,3}$ is also infinite.

[34] Townships with no mechanized firms are captured in the model by allowing both $\eta_{n,2}$ and $\eta_{n,3}$ to approach infinity, in which case $\varphi_{n,2}$ and $\varphi_{n,3}$ are both infinite.

(16) implies:

$$\varphi_{n,2} = \frac{w_n}{\rho P}\left(\frac{\sigma f w_n\left(\eta_{n,2}-1\right)}{R\left(\gamma_2^{\sigma-1}-1\right)}\right)^{\frac{1}{\sigma-1}}$$

Active firms that are not productive enough to use water or steam choose to operate without mechanized power. Define $\varphi_{n,1}$ as the operating cut-off, such that a non-mechanized firm with this productivity level is just indifferent between operating and exiting after learning its productivity type. By definition a firm that draws $\varphi_{n,1}$ makes zero profit, $\pi(\varphi_{n,1},1)=0$. From the profit function in equation (16) it follows that $\varphi_{n,1} = w_n\left(\sigma f w_n\right)^{\frac{1}{\sigma-1}} / \left(\rho P R^{\frac{1}{\sigma-1}}\right)$.

Given the profit function in equation (16), and the cut-off equations, the expected profit of an active firm is given by the following *zero-cut-off-profit* condition:

$$\frac{E(\pi_n|\varphi\geq\varphi_{n,1})}{w_n} = \frac{f(\sigma-1)}{\theta-\sigma+1}\left(1+\left(\eta_{n,2}-1\right)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}+\left(\eta_{n,3}-\eta_{n,2}\right)\left(\frac{\eta_{n,3}-\eta_{n,2}}{\gamma_3^{\sigma-1}-\gamma_2^{\sigma-1}}\right)^{-\frac{\theta}{\sigma-1}}\right) \quad (ZCP)$$

Finally, at stage A firms will enter in each township up until the point that expected profits are equal to the fixed cost of entry, $\left(1-G(\varphi_{n,1})\right)E(\pi_n|\varphi\geq\varphi_{n,1}) = w_n f_e$. The equilibrium operating cut-off, $\varphi_{n,1}$, can be solved by combining this free entry condition with the ZCP condition:

$$\varphi_{n,1} = \left(\frac{f(\sigma-1)}{f_e(\theta-\sigma+1)}\left[1+\left(\eta_{n,2}-1\right)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}+\left(\eta_{n,3}-\eta_{n,2}\right)\left(\frac{\eta_{n,3}-\eta_{n,2}}{\gamma_3^{\sigma-1}-\gamma_2^{\sigma-1}}\right)^{-\frac{\theta}{\sigma-1}}\right]\right)^{\frac{1}{\theta}} \quad (17)$$

Free mobility households ensures that welfare is equalized across all townships. This equilibrium condition can be used to solve for the endogenous population of each township, L_n. Define $\mathbb{W}_n$ as the welfare of a representative household in township n. From the utility function in equation (1), welfare can be written as function the township wage, w_n, the CES price index, P, and population density, L_n/H_n:

$$\mathbb{W}_n = \left(\frac{w_n}{P}\right)^{\alpha}\left(\frac{H_n}{L_n}\right)^{1-\alpha} = \left(\frac{R}{\sigma f}\rho^{\sigma-1}\right)^{\frac{\alpha}{\sigma-1}}\left(\frac{\varphi_{n,1}}{w_n^{\frac{1}{\sigma-1}}}\right)^{\alpha}\left(\frac{H_n}{L_n}\right)^{1-\alpha} = \mathbb{W} \quad (18)$$

The second equality follows from the revenue function in equation (16), evaluated at the operating cutoff $\varphi_{n,1}$. The final equality holds by the free mobility of households. For any two townships i and n, equation (18) implies:

$$\frac{w_n}{w_i} = \left(\frac{\varphi_{n,1}}{\varphi_{i,1}}\right)^{\sigma-1}\left(\frac{L_i H_n}{L_n H_i}\right)^{\frac{(1-\alpha)(\sigma-1)}{\alpha}} \quad (19)$$

A second equation for the relative wage between townships i and n can be obtained from the condition that a firm that draws $\varphi_{n,1}$ makes zero profit in each township. This condition along with the definition of CES demand implies that $w_n/w_i = \left(\varphi_{n,1}/\varphi_{i,1}\right)^{(\sigma-1)/\sigma}$. The equilibrium population in each township can be derived by combining this expression with equation (19), and using the aggregate labor supply condition $\sum_{n=1}^{N} L_n = L$:

$$L_n = L\frac{H_n\varphi_{n,1}^{\frac{\sigma-1}{\sigma}\frac{\alpha}{1-\alpha}}}{\sum_{i=1}^{N} H_i\varphi_{i,1}^{\frac{\sigma-1}{\sigma}\frac{\alpha}{1-\alpha}}} \quad (20)$$

Finally, the township labor market clearing condition is used to solve for the equilibrium measure of firms, M_n, which yields the following expression:

$$M_n = \frac{\sigma - 1}{\sigma \theta fe} \frac{L_n}{\varphi_{n,1}^{\theta}} \tag{21}$$

6.1.2 Deriving the Welfare and Urbanization Statistics

This appendix derives the welfare and urbanization statistics that are presented in section 3.2.

To derive the urbanization statistic, recall the representation of welfare in equation (18): $\mathbb{W} = (w_n/P)^{\alpha}(H_n/L_n)^{1-\alpha}$. Free mobility implies the following relative wage equation: $w_i/w_n = (L_i H_n/(H_i L_n))^{(1-\alpha)/\alpha}$. Note that this relative wage equation holds in both the observed and counterfactual economy, and recall that the land endowment of each township, H_n, is held constant in the counterfactual exercise. It follows that the relative wage growth for any two townships, i and n, is given by the following equation:

$$ln\left(\frac{w_i/w_i^{'}}{w_n/w_n^{'}}\right) = \frac{1-\alpha}{\alpha} ln\left(\frac{L_i/L_i^{'}}{L_n/L_n^{'}}\right) \tag{22}$$

Relative to the counterfactual economy, townships with higher wage growth also have higher population growth, such that the welfare gains from higher nominal wages are offset by increased congestion.

Recall from section (6.1.1), that the relative wage between townships i and n can also be described by the operating firm cut-off ratio, $w_i/w_n = (\varphi_{i,1}/\varphi_{n,1})^{(\sigma-1)/\sigma}$. The urbanization statistic is derived by substituting this expression in equation (22), and using that aggregate population condition, $\sum_{n=1}^{N} L_n = \sum_{n=1}^{N} L_n^{'} = L$, which yields:

$$ln\left(\frac{L_n}{L_n^{'}}\right) = \frac{\alpha(\sigma-1)}{(1-\alpha)\sigma} ln\left(\frac{\phi_{n1}}{\phi_{n1}^{'}}\right) + ln\left(\sum_{n=1}^{N} \frac{L_i}{L}\left(\frac{\phi_{i1}^{'}}{\phi_{i1}}\right)^{\frac{\alpha(\sigma-1)}{(1-\alpha)\sigma}}\right) \tag{23}$$

The representation of the urbanization statistic in equation (23) is expressed in terms of the model's parameters α and σ, aggregate values from the observed economy, L_n and L, and the operating cut-off ratio, $\phi_{n1}/\phi_{n1}^{'}$. I will now follow a similar approach in deriving the welfare statistic, $ln\left(\mathbb{W}^{'}/\mathbb{W}\right)$, and then derive an expression for the operating cut-off ratio that can be estimated from the data described in section 2.

To derive an expression for the welfare statistic note that the following zero profit condition holds in the observed and counterfactual economy: $\pi(\varphi_{n1}, 1) = \pi\left(\varphi_{n1}^{'}, 1\right) = 0$. This condition and the definition of CES demand implies that the first term in equation (7) is:[35]

$$\alpha ln\left(\frac{w_n P^{'}}{w_n^{'} P}\right) = \frac{\alpha}{\sigma-1} ln\left(\frac{R R_n^{'}}{R^{'} R_n}\right) + \frac{\alpha}{\sigma-1} ln\left(\frac{L_n}{L_n^{'}}\right) + \alpha ln\left(\phi_{n1}/\phi_{n1}^{'}\right), \tag{24}$$

where R_n is aggregate firm revenue in township n. Combining equations (7) and (24) yields the following

[35] This derivation also makes use of the result that wage income equals firm revenue in each township, $w_n L_n = R_n$.

welfare equation:[36]

$$ln\left(\mathbb{W}'/\mathbb{W}\right) = \frac{\alpha}{1-\sigma} ln\left(\sum_{n=1}^{N} \frac{R_n}{R}\left(\frac{\varphi_{n1}}{\varphi'_{n1}}\right)^{1-\sigma}\left(\frac{L_n}{L'_n}\right)^{\frac{\sigma(1-\alpha)-1}{\alpha}}\right) \tag{25}$$

The final step is to derive a convenient expression for the operating cut-off ratio, ϕ_{n1}/ϕ'_{n1}, such that the urbanization and welfare statistics, given by (23) and (25) respectively, can be estimated from the data described in section 2 . From equations (4) and (5), the operating cut-off ratio is:

$$\frac{\phi_{n1}}{\phi'_{n1}} = \left(\frac{1+(\eta_{n,2}-1)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}} + (\eta_{n,3}-\eta_{n,2})\left(\frac{\eta_{n,3}-\eta_{n,2}}{\gamma_2^{\sigma-1}-\gamma_3^{\sigma-1}}\right)^{-\frac{\theta}{\sigma-1}}}{1+(\eta_{n,2}-1)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}}\right)^{\frac{1}{\theta}} \tag{26}$$

The term in parenthesis in the numerator can be written:

$$1+(\eta_{n,2}-1)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}} + (\eta_{n,3}-\eta_{n,2})\left(\frac{\eta_{n,3}-\eta_{n,2}}{\gamma_2^{\sigma-1}-\gamma_3^{\sigma-1}}\right)^{-\frac{\theta}{\sigma-1}} = \frac{\theta-\sigma+1}{\theta\sigma f}\frac{L_n}{M_n} = \frac{(\theta-\sigma+1)\,R_n}{\theta\sigma f w_n M_n}$$

The first equality follows from the definition of the equilibrium measure of firms in township n, equation (21), and the second equality makes use of the result that $R_n = w_n L_n$. Combining this expression with equation (26) yields:

$$\frac{\phi_{n1}}{\phi'_{n1}} = \left(\frac{R_n}{\frac{\theta}{\theta-\sigma+1}\sigma f w_n M_n\left(1+(\eta_{n,2}-1)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}\right)}\right)^{\frac{1}{\theta}} = \left(1+\frac{R_{n,3}}{R_n}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma}-1\right)\right)^{-\frac{1}{\theta}}, \tag{27}$$

where $R_{n,3}$ is the revenue of steam powered firms in township n. The final equality in equation (27) makes use of the following two results:

$$R_n = \frac{\theta}{\theta-\sigma+1}\sigma f w_n M_n\left(1+(\eta_{n,2}-1)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}} + (\eta_{n,3}-\eta_{n,2})\left(\frac{\eta_{n,3}-\eta_{n,2}}{\gamma_3^{\sigma-1}-\gamma_2^{\sigma-1}}\right)^{-\frac{\theta}{\sigma-1}}\right),$$

$$R_{n,3}\frac{\gamma_3^{\sigma-1}-\gamma_2^{\sigma-1}}{\gamma_3^{\sigma-1}} = \frac{\theta}{\theta-\sigma+1}\sigma f w_n M_n(\eta_{n,3}-\eta_{n,2})\left(\frac{\eta_{n,3}-\eta_{n,2}}{\gamma_3^{\sigma-1}-\gamma_2^{\sigma-1}}\right)^{-\frac{\theta}{\sigma-1}}$$

Note that for townships without steam powered firms, $\phi_{n1}/\phi'_{n1} = 1$, by equation (26). This is consistent with operating cut-off ratio given by equation (27), since in any such township $R_{n,3} = 0$, which implies $\phi_{n1}/\phi'_{n1} = 1$.

Combining equation (27) with the equations (23) and (25), yields the following welfare and urbanization

[36]To derive this equation, re-arrange and sum over R'_n/R' for all $n \in \{1, 2, ..., N\}$, and note that this sum equals one.

statistics:

$$ln\left(\mathbb{W}'/\mathbb{W}\right)=\frac{\alpha}{1-\sigma}ln\left(\sum_{n=1}^{N}\frac{R_n}{R}\left(1+\frac{R_{n,3}}{R_n}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma}-1\right)\right)^{\frac{\sigma-1}{\theta}}\left(\frac{L_n}{L_n'}\right)^{\frac{\sigma(1-\alpha)-1}{\alpha}}\right) \tag{28}$$

$$ln\left(\frac{L_n}{L_n'}\right)=\frac{\alpha\left(1-\sigma\right)}{\theta\left(1-\alpha\right)\sigma}ln\left(1+\frac{R_{n,3}}{R_n}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma}-1\right)\right)+ln\left(\sum_{n=1}^{N}\frac{L_i}{L}\left(1+\frac{R_{i,3}}{R_i}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma}-1\right)\right)^{\frac{\alpha(\sigma-1)}{\theta(1-\alpha)\sigma}}\right) \tag{29}$$

6.1.3 Townships with Steam Powered Firms, but No Water Powered Firms

In the model developed in section 3.1, I assumed that all townships having steam powered firms also have water powered firms. This assumption was made in the interest of deriving welfare and urbanization statistics that can be tractably estimated, given the data described in section 2. This appendix relaxes this assumption, and derives the welfare and urbanization statistics under this generalization of the model.[37] Throughout this appendix, I denote equilibrium values in this more general setting with asterisks. For example, the operating cut-off ratio in the generalized model is denoted $\phi_{n1}^*/\phi_{n1}^{'*}$

I begin by deriving the parameter restrictions that are consistent with a township having steam, but not water powered firms. For such a township, define $\varphi_{n,3}^*$ as the cut-off, such that a firm with this productivity level is just indifferent with respect to using steam power and operating without mechanized power. That is, $\pi_n(\varphi_{n,3}^*,\gamma_3)=\pi_n(\varphi_{n,3}^*,1)$, which from the profit function in equation (16) implies:

$$\varphi_{n,3}^*=\frac{w_n^*}{\rho P^*}\left(\frac{\sigma f w_n\left(\eta_{n,3}-1\right)}{R^*\left(\gamma_3^{\sigma-1}-1\right)}\right)^{\frac{1}{\sigma-1}} \tag{30}$$

For there to be no water powered firms, it must be the case that it would not be profitable for any non-mechanized firm in the township to switch to water power. The assumption that $\gamma_2>1$ implies that the profit differential between water and non-mechanized firms, $\pi_n(\varphi,\gamma_2)-\pi_n(\varphi,1)$, is strictly increasing in φ. Therefore, if $\pi_n(\varphi_{n,3}^*,\gamma_2)<\pi_n(\varphi_{n,3}^*,1)$, then $\pi_n(\varphi,\gamma_2)<\pi_n(\varphi,1)$ for all non-mechanized firms, $\varphi\in\left[\varphi_{n,1}^*,\varphi_{n,3}^*\right]$. From the definition of the profit function in equation (16), it follows that in any township having steam but not water powered firms, the following inequality must hold:

$$\frac{R^*P^{*\sigma-1}}{w_n^*\sigma f}\left(\frac{\rho\varphi_{n,3}^*}{w_n^*}\right)^{\sigma-1}<\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\qquad\Longrightarrow\qquad\frac{\eta_{n,3}-1}{\gamma_3^{\sigma-1}-1}<\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1} \tag{31}$$

The rightmost inequality in equation (31) is a parameter restriction that must hold for any township with steam but no water powered firms. I now show that for any such township, this parameter restriction implies that the operating cut-off ratio, $\phi_{n1}^*/\phi_{n1}^{'*}$, is greater than its corresponding value in the restricted framework, $\phi_{n1}/\phi_{n1}^{'}$, as derived in section 6.1.2. For such a township, the operating cut-off ratio in the

[37] The assumption that there are non-mechanized firms in each township is maintained in this appendix. As noted in section 3.1, this assumption is consistent with the data since over 95 percent of Ontario townships in 1871 had non-mechanized firms.

generalized model is:

$$\frac{\phi^*_{n1}}{\phi'^*_{n1}} = \left(\frac{1 + (\eta_{n,3} - 1)\left(\frac{\eta_{n,3}-1}{\gamma_3^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}}{1 + (\eta_{n,2} - 1)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}} \right)^{\frac{1}{\theta}} \tag{32}$$

Compare this value to the cut-off ratio that was derived in section 6.1.2, ϕ_{n1}/ϕ'_{n1}, as given by equation (27)[38]:

$$\phi_{n1}/\phi'_{n1} = \left(1 + \frac{R_{n,3}}{R_n}\left(\left(\frac{\gamma_3}{\gamma_2}\right)^{1-\sigma} - 1\right)\right)^{-\frac{1}{\theta}} = \left(\frac{1 + (\eta_{n,3} - 1)\left(\frac{\eta_{n,3}-1}{\gamma_3^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}}{1 + \frac{\gamma_2^{\sigma-1}-1}{\gamma_3^{\sigma-1}-1}(\eta_{n,3} - 1)\left(\frac{\eta_{n,3}-1}{\gamma_3^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}} \right)^{\frac{1}{\theta}} \tag{33}$$

Note that the operating cut-off ratio in the generalized model, ϕ^*_{n1}/ϕ'^*_{n1}, is greater than its corresponding value in the restricted model, ϕ_{n1}/ϕ'_{n1}:

$$\left(\frac{1 + (\eta_{n,3} - 1)\left(\frac{\eta_{n,3}-1}{\gamma_3^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}}{1 + (\eta_{n,2} - 1)\left(\frac{\eta_{n,2}-1}{\gamma_2^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}} \right)^{\frac{1}{\theta}} > \left(\frac{1 + (\eta_{n,3} - 1)\left(\frac{\eta_{n,3}-1}{\gamma_3^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}}{1 + \frac{\gamma_2^{\sigma-1}-1}{\gamma_3^{\sigma-1}-1}(\eta_{n,3} - 1)\left(\frac{\eta_{n,3}-1}{\gamma_3^{\sigma-1}-1}\right)^{-\frac{\theta}{\sigma-1}}} \right)^{\frac{1}{\theta}}$$

$$1 + \frac{\gamma_2^{\sigma-1} - 1}{\gamma_3^{\sigma-1} - 1}(\eta_{n,3} - 1)\left(\frac{\eta_{n,3} - 1}{\gamma_3^{\sigma-1} - 1}\right)^{-\frac{\theta}{\sigma-1}} > 1 + (\eta_{n,2} - 1)\left(\frac{\eta_{n,2} - 1}{\gamma_2^{\sigma-1} - 1}\right)^{-\frac{\theta}{\sigma-1}}$$

$$\frac{\eta_{n,3} - 1}{\gamma_3^{\sigma-1} - 1} < \frac{\eta_{n,2} - 1}{\gamma_2^{\sigma-1} - 1} \tag{34}$$

Where the inequality in the final line of equation (34) follows from the parameter restriction that were derived in equation (31). To summarize, equation (32) shows the operating cut-off ratio for townships with steam but no water powered firms. Equation (34) establishes that this cut-off ratio is greater than its corresponding value in the restricted model that excludes such townships by assumption.

The motivation for restricting the model to exclude townships with steam but no water powered firms was to retain tractability in estimating the welfare and urbanization statistics. I complete this section by discussing how this simplifying assumption affects the urbanization and welfare statistics that were derived in section 6.1.2.

Note from equation (23) that the urbanization statistic is strictly increasing in each township's own operating cut-off:

$$\frac{\partial\left(ln\left(L_n/L'_n\right)\right)}{\partial\left(\varphi_{n,1}/\varphi'_{n,1}\right)} = \frac{\alpha(\sigma-1)}{(1-\alpha)\sigma}\frac{\varphi'_{n,1}}{\varphi_{n,1}}\left(\sum_{i\in N, i\neq n}\frac{L_i}{L}\left(\frac{\varphi_{i1}}{\varphi'_{i1}}\right)^{\frac{\alpha(1-\sigma)}{(1-\alpha)\sigma}}\right)\left(\sum_{i=1}^{N}\frac{L_i}{L}\left(\frac{\varphi_{i1}}{\varphi'_{i1}}\right)^{\frac{\alpha(1-\sigma)}{(1-\alpha)\sigma}}\right)^{-1} > 0 \tag{35}$$

[38]Equation (33) makes use of the derived values of R_n and $R_{n,3}$ for a township with steam but no water powered firms:

$$R_n = \frac{\theta\sigma f w_n M_n}{\theta - \sigma + 1}\left(1 + (\eta_{n,3} - 1)\left(\frac{\eta_{n,3} - 1}{\gamma_3^{\sigma-1} - 1}\right)^{-\frac{\theta}{\sigma-1}}\right)^{-\frac{\theta}{\sigma-1}}, \quad R_{n,3} = \frac{\theta\sigma f w_n M_n}{\theta - \sigma + 1}(\eta_{n,3} - 1)\left(\frac{\eta_{n,3} - 1}{\gamma_3^{\sigma-1} - 1}\right)^{-\frac{\theta}{\sigma-1}}\frac{\gamma_3^{\sigma-1}}{\gamma_3^{\sigma-1} - 1}$$

Conversely, township n's the urbanization statistic is strictly decreasing in any other township i's operating cut-off ratio:

$$\frac{\partial\left(ln\left(L_n/L_n^{'}\right)\right)}{\partial\left(\varphi_{i,1}/\varphi_{i,1}^{'}\right)} = -\frac{\alpha(\sigma-1)}{(1-\alpha)\sigma}\left(\frac{L_i}{L}\left(\frac{\varphi_{i1}}{\varphi_{i1}^{'}}\right)^{\frac{\alpha(1-\sigma)}{(1-\alpha)\sigma}-1}\right)\left(\sum_{i=1}^{N}\frac{L_i}{L}\left(\frac{\varphi_{i1}}{\varphi_{i1}^{'}}\right)^{\frac{\alpha(1-\sigma)}{(1-\alpha)\sigma}}\right)^{-1} < 0 \tag{36}$$

Two implications can be drawn from the partial derivatives in equations (35) and (36). First, for any township with steam but no water powered firms, the urbanization statistic given by equation (9) will typically be a lower bound, relative to its value in the more general setting.[39] Second, for townships of any other type, equation (9) is an upper bound, so long as there is at least one townships in the data with steam but no water powered firms. The follows directly from the partial derivative in equation (36).

The following partial derivative evaluates the change in the welfare statistic from an increase in any township n's operating cut-off ratio:

$$\frac{\partial\left(ln\left(\mathbb{W}^{'}/\mathbb{W}\right)\right)}{\partial\left(\varphi_{n,1}/\varphi_{n,1}^{'}\right)} = \frac{\alpha(\sigma(1-\alpha)-1)}{\sigma(1-\alpha)}\frac{L_n}{L}\left(\frac{\varphi_{n,1}}{\varphi_{n,1}^{'}}\right)^{\frac{\alpha(1-\sigma)}{(1-\alpha)\sigma}-1}\left(\sum_{n=1}^{N}\frac{L_n}{L}\left(\frac{\varphi_{n1}}{\varphi_{n1}^{'}}\right)^{\frac{\alpha(1-\sigma)}{(1-\alpha)\sigma}}\right)^{-1}$$

$$+\frac{\alpha}{\sigma(1-\alpha)}\frac{R_n}{R}\left(\frac{\varphi_{n,1}}{\varphi_{n,1}^{'}}\right)^{\frac{(1-\sigma)}{(1-\alpha)\sigma}-1}\left(\sum_{n=1}^{N}\frac{R_n}{R}\left(\frac{\varphi_{n1}}{\varphi_{n1}^{'}}\right)^{\frac{(1-\sigma)}{(1-\alpha)\sigma}}\right)^{-1}$$

The sign of this partial derivative is indeterminate, as the term $\sigma(1-\alpha)-1$ may be positive or negative. This implies that restricting the model to exclude townships with steam but not water powered firms has an ambiguous effect on the estimated welfare gains from steam.

6.2 Empirical Appendix

6.2.1 Auxiliary Regressions used in Instrumental Variables Estimation

[39] The reason that this is typically, but not necessarily the case, is because there may be more than one township with steam but no water powered firms. If this is the case, then the downward bias on township n's urbanization statistic resulting from the fact that $\phi_{n1}^{*}/\phi_{n1}^{'*} > \phi_{n1}/\phi_{n1}^{'}$, could be less than the upward bias that results from $\phi_{i1}^{*}/\phi_{i1}^{'*} > \phi_{i1}/\phi_{i1}^{'}$, for all $i \neq n$, such that i and n are townships with steam but no water powered firms.

Table 5: Linear Regression of Steam Power on Instrumental Variable

Variable	(1)	(2)
$ln(slope_n)$	-0.209***	-0.199***
	(0.0237)	(0.0247)
$ln(Population_n)$	0.128***	
	(0.0208)	
$ln(Land\ Area_n)$	-0.0719***	
	(0.0104)	
$ln(PopulationDensity_n)$		0.0800***
		(0.0121)
$ln(M_{n,mech}/M_n)$	-0.0392	-0.0421
	(0.0287)	(0.0294)
SIC Industry Controls	Yes	Yes
Observations	4,097	4,097
R-squared	0.223	0.217
F-Statistic	77.80	64.63

Standard errors are clustered at the township level. *** p<0.01, ** p<0.05, * p<0.1. Coefficient estimates for industry indicator variables are not reported but are available upon request.

6.2.2 Sensitivity Analysis: the Elasticity of Substitution, σ

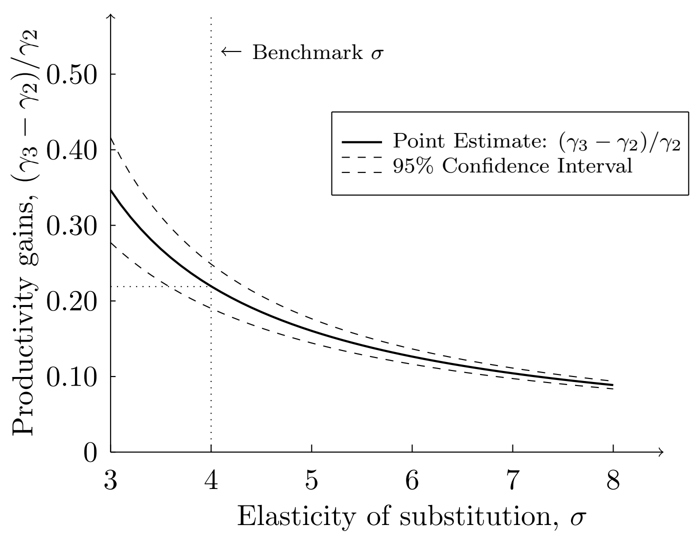


Figure 4: Firm-Level Productivity Gains from Steam Power Adoption

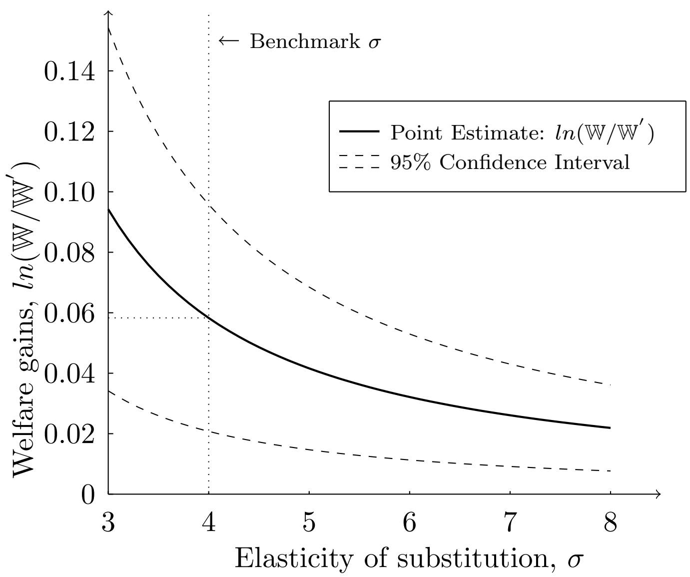


Figure 5: Aggregate Welfare Gains from Steam Power Adoption