Firm Heterogeneity and the Gains from Technological Adoption: Theory and Measurement

Alex W. Chernoff*

January 2, 2016

Abstract

This paper develops a tractable framework for estimating the productivity and welfare gains from the introduction of a general purpose technology that explicitly accounts for the effects of firm heterogeneity in adoption. I develop a theoretical model where monopolistically competitive firms are heterogeneous with respect to their productivity and choose to operate using one of several different technologies. The theoretical model is used to derive a statistic that summarizes the welfare gains from the introduction and adoption of a new technology. I consider two applications of the theoretical framework: steam power in the nineteenth century, and internet-enabled mobile devices (i.e. smartphones and tablets) in the twenty-first century. For steam power, I estimate the welfare statistic using nineteenth century firm-level data on steam power adoption in the Canadian manufacturing sector. I exploit exogenous variation in geography to estimate several structural parameters of the model. My results indicate that the use of steam power resulted in a 15.1 percent increase in firm-level productivity and a 3.0-5.2 percent increase in welfare. The second empirical application uses firm-level data on the adoption of mobile devices in the Canadian private sector over the period 2012-2013. I show that mobile device adoption is positively correlated with firm size and productivity, which is consistent with the predictions of the theoretical model. My current research is focused on estimating the welfare and productivity gains from mobile device adoption, and comparing these results with my estimates of the gains from steam power adoption.

JEL classification: O14, O47, N61

Keywords: Firm Heterogeneity, Productivity, General Purpose Technology, Steam, Information and Communications Technologies

---

*Queen's University, Ontario, Canada, Email: chernoff@econ.queensu.ca. I would like to thank my PhD supervisors Beverly Lapham, Huw Lloyd-Ellis, and Ian Keay. 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 gratefully acknowledge financial support provided by the Social Sciences and Humanities Research Council of Canada, the Network to Study Productivity in Canada from a Firm-Level Perspective, and Industry Canada.

1 Introduction

There are several theoretical and empirical challenges associated with estimating the aggregate effects of technological adoption. Aggregate models of technological adoption abstract from important questions regarding why only some firms adopt certain technologies and how this heterogeneity affects aggregate outcomes. While models of firm heterogeneity are increasingly popular, tractability in aggregation poses a challenge to estimating the macroeconomic effects of technological adoption in micro-founded models. From an empirical perspective, the inherently endogenous relationship between productivity and technological adoption raises issues of identification that are not easily resolved. This paper develops and estimates a model of firm heterogeneity and technological adoption that overcomes these challenges.

I develop a model of technological adoption with monopolistically competitive firms that are heterogeneous with respect to their productivity. Building on the framework of Bustos (2011), firms make a discrete choice to adopt one of several different technologies that vary in terms of productivity and their fixed cost of adoption. Despite the superiority of the most productive technology, firm-level heterogeneity and differences in the cost of adoption result in each of the various technologies being used by some fraction of firms in equilibrium. I use the theoretical model to develop a welfare statistic that measures the welfare gains from the introduction of a new technology that is superior to existing technologies. Intuitively, the welfare statistic depends on the counterfactual revenue of firms using the new technology, had they been forced to operate instead with the second most productive technology.

I consider two empirical applications of the theoretical framework: steam power in the late nineteenth century, and internet-enabled mobile devices in the twenty-first century. Studying steam power adoption is a well-suited application of the model for a number of reasons. As a general purpose technology (GPT), my analysis provides insight into the aggregate effects of GPTs when there is firm heterogeneity in adoption.[1] Studying steam-power adoption in Canada is representative of the second generation of countries that entered into industrialization during the late nineteenth

[1] Other papers that have analyzed steam power as a GPT include Crafts (2004) and Rosenberg and Trajtenberg (2004).

century, such as Argentina, Australia, and Sweden. The data used in this paper covers the universe of firms in the Canadian manufacturing sector in 1871. Given the objectives of this paper, the universal nature of the data renders it superior to comparable data for the U.S., such as the Atack and Bateman (1999) samples from the nineteenth century U.S. Manufacturing Censuses. This feature of the data ensures that the aggregate welfare estimates are representative, and provides an adequate sample size for consistent estimation of the effect of steam-power adoption on firm-level productivity.

As noted at the onset, an empirical challenge associated with estimating the effect of technological adoption relates to identification. My identification strategy draws on the theory that variation in agricultural land endowments partially explain differential rates of industrialization in the nineteenth century, as has been argued by Rothbarth (1946), Habakkuk (1962), Clarke and Summers (1980), and McCallum (1980). Using an instrumental variables identification strategy, I estimate that steam power adoption increased firm-level labor productivity by approximately 15.1 percent relative to water power. To calculate the welfare statistic, I estimate two structural equations from the theoretical model. My results indicate that steam power adoption resulted in a 3.0-5.2 percent increase in welfare. The empirical framework developed in this paper draws from methodology that has been used to estimate the welfare gains from international trade in models with heterogeneous firms. Comparing my results to this literature, I show that the estimated welfare gains from steam power adoption are approximately one and a half times as large as contemporary estimates of the welfare gains from trade.

The second empirical application uses newly available enterprise-level data on the adoption of internet-enabled mobile devices in the Canadian private sector over the period 2012-2013. The empirical results I present on mobile device adoption are preliminary and are used to motivate my ongoing research on this topic. I show that the enterprise-level adoption of mobile devices is positively correlated with enterprise-scale and labor productivity. These results parallel the characteristics of firms that adopted steam power in the late nineteenth century, and are also consistent with the theoretical model developed in this paper. The objective of my current research is to use the theoretical framework of this paper to estimate the productivity and welfare gains from mobile devices. I will compare the magnitude of these estimates to the productivity and welfare gains from steam power, while acknowledging the limitations of comparing GPTs from different eras in the Canadian economy.

This paper contributes to the literature on GPTs by developing a micro-founded empirical framework for estimating the aggregate welfare gains from the introduction of a GPT when firm-heterogeneity factors heavily in the pattern of adoption. My theoretical framework extends Bustos's (2011) model of technological choice by allowing for multiple technologies and by deriving a welfare statistic that can be estimated using firm-level data. This paper also contributes to the literature on steam power, by quantifying the effect of adoption on firm-level productivity and aggregate welfare. There is an extensive literature that documents the importance of steam power to growth and industrialization during the nineteenth century.[2] My work is similar to Crafts (2004), who estimates the aggregate effect of steam power on economic growth in the British economy over the period 1760-1910. Relative to information and communications technologies (ICTs), Crafts (2004) finds that that steam technology resulted in only modest contributions to labor productivity growth, peaking during the period 1850-1870 at 0.41 percent per year. However, Crafts (2004) uses a growth accounting framework which ignores the effects of firm heterogeneity. My results suggest that gains stemming from firm heterogeneity in the pattern of steam power adoption were important, as the use of steam power was highly correlated with firm size and productivity.

The remainder of the paper is organized as follows: Section 2 describes the data on steam power adoption, Section 3 presents the theoretical model, Section 4 presents the empirical application of the model to steam power, Section 5 describes the data on internet-enabled mobile devices, and Section 6 concludes.

[2] 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).

2 Firm Heterogeneity in GPT Adoption: the Case of Steam Power

An objective of this paper is to develop a tractable framework for estimating the gains from the introduction of a GPT that explicitly accounts for the effects of firm heterogeneity in adoption. The application of this framework in Section 4 considers the case of steam power in the Canadian manufacturing sector in the late nineteenth century. This section presents a number of empirical results regarding the use of steam power in Canadian manufacturing during this era that guide the development of the theoretical and empirical frameworks that are used in the remainder of the paper.

This paper uses data from the Canadian Industry in 1871 Project (CANIND71), which is a digitized record of the universe of manufacturing establishments that were enumerated in the 1871 Manufacturing Census of Canada.[3] Firm-level data on steam power adoption in the nineteenth century are available for other countries, notably including the Atack and Bateman (1999) samples from the nineteenth century U.S. Manufacturing Censuses. However given the objectives of this paper, the universal coverage in the CANIND71 database renders these data superior to the Atack and Bateman (1999) samples on account of two factors. Firstly, the empirical methodologies that are

[3]The history of the CANIND71 database is documented in Bloomfield and Bloomfield (1989). While the CANIND71 database is comprehensive, there are measurement issues in the raw manuscript data that must be reconciled prior to analysis. In particular, Inwood (1995) notes that multi-process establishments were occasionally decomposed by enumerators into multiple manuscript entries based on the distinct industrial activities carried out by the firm. Inwood (1995) finds that approximately 10 percent of the census manuscript entries were establishments that had been decomposed by enumerators. Using the data available in the CANIND71 database, I follow Inwood's (1995) procedure for identifying and reconstituting multi-process 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. Reconstitution of the CANIND71 database reduces the number of establishments from 45,070 to 43,098.

developed in Section 4 require estimation from various sub-samples, such as the sub-sample of firms using steam power. The large sub-samples that are available through the CANIND71 database ensure that the asymptotic properties of the estimators that are used can be reliably invoked. Secondly, as the main objective of this paper is to measure the aggregate welfare effect of steam power adoption, the universal coverage in the CANIND71 database minimizes the vulnerability of the results to sampling variation in estimation.

For the purposes of my analysis I exclude 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 SIC major industry classes: agricultural services, forestry, mining, construction, gas and water utilities, personal and business services, and trade (including repair services). Excluding these industries reduces the sample size to 36,902. A further reduction in the sample occurs from excluding 62 firms that were located in remote census districts.[4] Finally, I exclude firms that did not report a positive value of value added, labor, or fixed capital, which results in a cleaned sample of 32,581 observations.

Table 1 summarizes firms' production characteristics across the different types of power used in Canadian manufacturing in 1871. The observation counts highlight the fact that only 8 percent of Canadian manufacturing establishments had adopted steam power in 1871.[5] By comparison, Atack et al. (2008) estimate that 20.1 percent of U.S. manufacturing establishments used steam power in 1870. However, caution is warranted in this comparison as differences in the enumeration practices between the Canadian and the U.S. censuses overstate the magnitude of this difference. In particular, 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 11 percent.[6]

[4]Remote census districts were identified as those having fewer than ten manufacturing establishments. These census districts were all located in remote northern regions of Ontario and Quebec.

[5]The classification of firms using steam power include 137 firms that reported using both steam and water power. All other firms in the sample were recorded as using only one of the five different types of power.

[6]For the purposes of converting the U.S. revenue threshold to Canadian dollars I use the official currency con-

Table 1: Production Profile by Type of Power Used

	Hand	Water	Steam	Animal	Wind	Full Sample
Employees	3.20	6.03	20.59	4.39	1.62	5.19
	(0.0594)	(0.271)	(0.933)	(0.186)	(0.154)	(0.104)
Fixed capital	711	3,736	11,534	1,154	591	2,189
	(21.77)	(165.5)	(575.3)	(67.80)	(98.25)	(60.97)
Revenue	2,809	9,801	30,361	3,138	1,183	6,409
	(70.79)	(419.6)	(1,773)	(200.0)	(243.5)	(177.4)
Value Added	1,360	3,098	14,279	1,840	399	2,755
	(30.86)	(142.8)	(803.0)	(99.75)	(56.02)	(76.30)
Observations	22,617	6,379	2,630	905	50	32,581

Standard errors for the mean estimates are in parentheses. Fixed capital, revenue, and value added are in nominal 1871 Canadian dollars.

The most striking pattern in Table 1 is the positive correlation between firm size and the use of steam power. The average number of workers employed by a firm using steam power was roughly four times that of the overall average for the manufacturing sector. Furthermore, the value of fixed capital, revenue, and value added were approximately five times higher for firms with steam power. These findings are consistent with Atack et al. (2008), who also find a positive relationship between firm size and steam power adoption in the U.S. manufacturing sector during this era. The scale dominance of steam powered firms is illustrated dramatically in Figure 1. This figure plots the empirical cumulative distribution functions (CDFs) for the three most common types of power used by Canadian manufacturing firms in 1871: hand, water, and steam. The percentage of firms operating below any level of output is always highest for hand, followed by water, and then steam powered firms. This ordering of the three main technologies in the distribution of firm output is replicated in the theoretical model developed in Section 3.

The final 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) regressions of labor productivity on steam and water power adoption. The dependent variable is the natural logarithm of labor productivity (value added divided by employment weighted by

version 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).

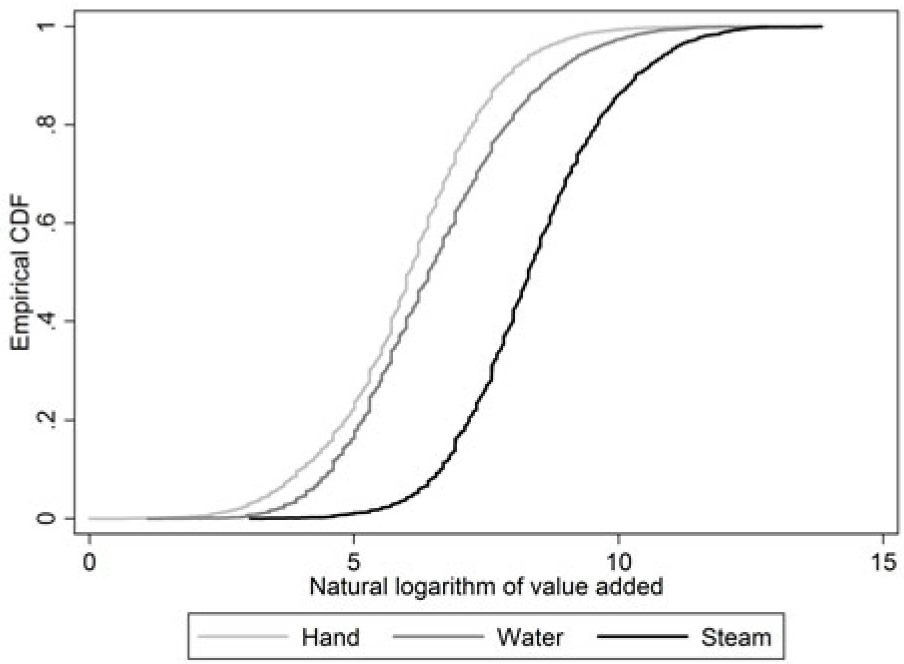

Figure 1: Empirical CDF of Natural Logarithm of Value Added by Type of Power Used

months in operation). Value added is measured as gross sales minus the value of raw material inputs. 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, the population in the firm's census district, and indicator variables for the firm's province of residence. Finally, column (3) includes an additional control variable for capital-intensity. Capital intensity is measured by the capital-labor ratio, and is calculated as the natural logarithm of fixed capital divided by employment weighted by months in operation.

The regression results in Table 2 indicate that steam and water power were correlated with higher labor productivity in the Canadian manufacturing sector in 1871, holding constant the other control variables in the regression analysis. The coefficient on the steam power variable is positive, and statistically significant at the 1 percent level in each specification. The coefficient on the water power variable is positive, and statistically significant at the 1 percent level in specifications (1) and (2), and positive and statistically significant at the 10 percent level in specification (3). The magnitude of water and steam power coefficients drops considerably in specification (3). This is

Table 2: Regression Analysis of Labor Productivity for Canadian Manufacturing in 1871

Variable	(1)	(2)	(3)
Steam Power	0.636***	0.489***	0.118***
	(0.0163)	(0.0176)	(0.0174)
Water Power	0.526***	0.490***	0.0256*
	(0.0134)	(0.0136)	(0.0147)
Five or Fewer Employees		-0.0939***	-0.131***
		(0.0126)	(0.0119)
Male Employees/Employees		0.614***	0.466***
		(0.0201)	(0.0188)
ln(District Population)		0.159***	0.109***
		(0.00832)	(0.00780)
ln(Capital/Labor)			0.252***
			(0.00421)
Province Controls	No	Yes	Yes
SIC Industry Controls	Yes	Yes	Yes
Observations	32,581	32,581	32,581
R-squared	0.158	0.206	0.306

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Coefficient estimates for province and industry indicator variables are not reported but are available upon request.

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. Atack et al. (2008) note that relative to water power, steam was a flexible power source that had the clear advantage of not being tied to a geographic location. This logic provides a reasonable explanation for the finding that steam is more highly correlated with productivity, after controlling for capital intensity.

The coefficients for the other control variables in Table 2 are consistent with the literature on the North American manufacturing sector during this era.[7] 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 Canadian 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

[7] 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 1 percent level in specifications (2) and (3).

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 census district 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 empirical analysis in this section can be summarized by three main results. First, a relatively small fraction of firms used steam power in the Canadian manufacturing sector in 1871. Second, the firms using steam power were large, with value added that was approximately five times the average for the Canadian manufacturing sector. Third, steam and water power adoption were both positively correlated with labor productivity. For both technologies this positive correlation persists when capital intensity is controlled for, however the magnitude and statistical significance of the results are strongest for steam.

The analysis in this section does not identify the causal effect of steam power adoption on productivity, and the analysis is also not capable of quantifying the effect of steam power on welfare in the aggregate economy. In Sections 3 and 4 I develop and estimate a model of heterogeneous firms and technological adoption to achieve these objectives.

3 Theory

This section develops a model of technological adoption with monopolistically competitive firms that are heterogeneous with respect to their productivity. The theoretical framework extends the model of Bustos (2011) by allowing for multiple technologies and by deriving a welfare statistic that can be estimated using firm-level data. In the model, firms make a discrete choice to operate using one of several different technologies. These technologies vary from lowest to highest productivity (output per worker). Despite the superiority of the most productive technology, firm-level heterogeneity and differences in the cost of adoption results in each of the production technologies being used by some fraction of firms in equilibrium. The ultimate objective of this section of the paper is to use the model to derive a statistic that measures the welfare gains from the introduction and adoption of the steam power. In Section 4, I illustrate how this welfare statistic can be estimated using firm-level data. For concreteness, the model below refers to firms' adoption of steam power, but the framework can be applied to estimate the gains from the adoption of other technologies.

3.1 Model

I consider an economy with a unit measure of identical households with CES utility:

$$U = \left(\int_{\omega \in \Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \quad \rho \equiv \frac{\sigma - 1}{\sigma}, \tag{1}$$

where $q(\omega)$ is the quantity of variety ω demanded by the representative household, σ is the elasticity of substitution between varieties, and Ω is the endogenous set of varieties produced. Each household has one unit of labor that is supplied inelastically to firms. Labor is defined as the numeraire, and the wage is normalized to one. Households also hold an equal share in the profits of producers.

Firms use the following technologies in production:

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

The parameter γ differentiates the production technology across the three different power sources. That is, the firms' choice to use hand, water, or steam power is formalized by the associated choice of production technology $q(\varphi, 1)$, $q(\varphi, \gamma_2)$, or $q(\varphi, \gamma_3)$ respectively. Steam power's superiority over water and hand power is captured in the assumption that $\gamma_3 > \gamma_2 > 1$, where I have normalized the productivity parameter for hand power to one for notational convenience.

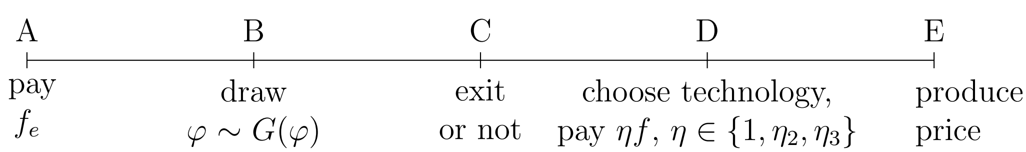

Figure 2: Firms' Timeline

The production functions in equation (2) also depend on the parameter φ, which is the firm's idiosyncratic productivity type. The firm's productivity type is realized through the process summarized in Figure 2. At stage A, a prospective firm chooses to enter the industry by paying a fixed cost f_e. All fixed costs are denoted in units of labor. After paying the entry cost, firms draw their productivity type, φ, from a Pareto CDF, $G(\varphi) = 1 - \varphi^{-\theta}$. 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. 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 type and pay a technology specific fixed cost of production. The firms' decision to use hand, water, or steam power is associated with a fixed cost of production equal to f, $\eta_2 f$, or $\eta_3 f$ respectively. I assume that $\eta_3 > \eta_2 > 1$ which implies that the cost of using steam is greater than for water, and that the cost of using water is greater than for hand power.

Solving the representative household's utility maximization problem yields the following CES demand function for variety ω:

$$q(\omega) = P^{\sigma-1} p(\omega)^{-\sigma}, \qquad P \equiv \left[\int_{\omega\in\Omega} p(\omega)^{1-\sigma} d\omega \right]^{\frac{1}{1-\sigma}} \tag{3}$$

where P is the CES price index.

Next I consider a firm's problem by working backward from stage E. Monopolistic competition and CES demand implies that the price chosen by a firm of type φ is: $p(\varphi,\gamma) = 1/(\rho\gamma\varphi)$. Combining the firm's pricing rule with household demand yields the following revenue and profit function:

$$r(\varphi,\gamma) = P^{\sigma-1} p(\varphi,\gamma)^{1-\sigma}, \qquad \pi(\varphi,\gamma) = \frac{r(\varphi,\gamma)}{\sigma} - \eta f \tag{4}$$

At stage D, firms choose a power type which affects profitability through the productivity parameter, γ, and through the fixed cost, $\eta \in \{1, \eta_2, \eta_3\}$. To match the empirical regularities presented in Section 2, I consider an equilibrium where all three technologies are used by some fraction of firms, and where steam is only used by the largest and most productive firms. Define φ_3 as the cut-off, such that a firm with this productivity level is just indifferent with respect to using steam or water power. That is, $\pi(\varphi_3, \gamma_3) = \pi(\varphi_3, \gamma_2)$, which from the profit function in equation (4) implies:

$$\varphi_3 = \frac{(\sigma f)^{\frac{1}{\sigma-1}}}{\rho P} \left[\frac{\eta_3 - \eta_2}{\gamma_3^{\sigma-1} - \gamma_2^{\sigma-1}} \right]^{\frac{1}{\sigma-1}} \tag{5}$$

Next consider the subset of firms choosing water power at stage D. Define φ_2 as the cut-off, such that a firm with this productivity level is just indifferent with respect to using water or hand power. That is, $\pi(\varphi_2, \gamma_2) = \pi(\varphi_2, 1)$, which from the profit function in equation (4) implies:

$$\varphi_2 = \frac{(\sigma f)^{\frac{1}{\sigma-1}}}{\rho P} \left(\frac{\eta_2 - 1}{\gamma_2^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}} \tag{6}$$

Active firms that are not productive enough to use water or steam choose to operate with hand power. Define φ_1 as the operating cut-off, such that a firm with this productivity level is just indifferent between operating with hand power and exiting after learning its productivity type. By definition a firm that draws φ_1 makes zero profit, $\pi(\varphi_1, 1) = 0$. From the profit function in equation (4) it follows that $\varphi_1 = (\sigma f)^{\frac{1}{\sigma-1}} / (\rho P)$. Combining this expression with equations (5) and (6) yields the following relative cut-off equations:[8]

$$\Delta_{31} \equiv \frac{\varphi_3}{\varphi_1} = \left(\frac{\eta_3 - \eta_2}{\gamma_3^{\sigma-1} - \gamma_2^{\sigma-1}} \right)^{\frac{1}{\sigma-1}}, \qquad \Delta_{21} \equiv \frac{\varphi_2}{\varphi_1} = \left(\frac{\eta_2 - 1}{\gamma_2^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}} \qquad (RC)$$

Define $\tilde{\pi}$ as the expected profit of an active firm, $\tilde{\pi} = E(\pi | \varphi \geq \varphi_1)$. Given the profit function in equation (4), and the relative cut-off equations (RC), the expected profit of an active firm can be derived and is given by the following *zero-cut-off-profit* condition:

$$\tilde{\pi} = \frac{f(\sigma - 1)}{\theta - \sigma + 1} \left(1 + (\eta_2 - 1)\Delta_{21}^{-\theta} + (\eta_3 - \eta_2)\Delta_{31}^{-\theta} \right) \qquad (ZCP)$$

Finally at stage A firms will enter the industry up until the point that expected profits are equal to the fixed cost of entry, $(1 - G(\varphi_1))\,\tilde{\pi} = f_e$. This free entry condition can be rewritten:

$$\tilde{\pi} = fe/\,(1 - G(\varphi_1)) = f_e \varphi_1^{\theta} \qquad (FE)$$

[8] In an equilibrium where all three technologies are used, the most productive firms use steam, and the least productive firms use hand power; it must be the case that $\varphi_1 < \varphi_2 < \varphi_3$. From equation (RC), it can be seen that this ordering of the cut-offs requires assuming that $\eta_2 - 1 > \gamma_2^{\sigma-1} - 1$ and $(\gamma_2^{\sigma-1} - 1)(\eta_3 - \eta_2) > (\gamma_3^{\sigma-1} - \gamma_2^{\sigma-1})(\eta_2 - 1)$. I impose these assumptions to derive an equilibrium that is consistent with the empirical regularities presented in Section 2.

The equilibrium operating cut-off, φ_1, can be solved by combining equations (ZCP) and (FE):

$$\varphi_1 = \frac{f(\sigma-1)}{fe(\theta-\sigma+1)}\left[1+(\eta_2-1)\Delta_{21}^{-\theta}+(\eta_3-\eta_2)\Delta_{31}^{-\theta}\right]^{\frac{1}{\theta}} \tag{7}$$

Define $\varphi_1^{'}$ as the equilibrium operating cut-off when steam power is not available in the economy. Solving the model without steam power yields:

$$\varphi_1^{'} = \frac{f(\sigma-1)}{fe(\theta-\sigma+1)}\left[1+(\eta_2-1)\Delta_{21}^{-\theta}\right]^{\frac{1}{\theta}} = \lim_{\eta_3\to\infty}\varphi_1 = \lim_{\gamma_3\to\gamma_2}\varphi_1 \tag{8}$$

Intuitively, when the cost of steam power adoption becomes prohibitively high, or when the productivity increase 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.

3.2 Measuring the Welfare Gains from Technological Adoption

In this section I derive a statistic that summarizes how the introduction and adoption of steam power affects welfare.[9] The primary objective of this exercise is to derive a representation of the welfare statistic that can be estimated from firm-level data. To derive this statistic, I compare the equilibrium of the model with hand, water, and steam power, to an equilibrium where only hand and water power are available to firms. Under CES preferences, welfare is measured by the real wage, which is equal to the reciprocal of the price index when the nominal wage is normalized to one, $\mathbb{W} = w/P = 1/P$. I denote welfare and other values in the equilibrium without steam with primes, $\mathbb{W}^{'} = 1/P^{'}$.

Proposition 1: The change in welfare from the introduction and adoption of steam power can be measured by the increase in the equilibrium operating cut-off, φ_1. That is:

$$\mathbb{W}/\mathbb{W}^{'} = P^{'}/P = \varphi_1/\varphi_1^{'} = \left(\frac{1+(\eta_2-1)\Delta_{21}^{-\theta}+(\eta_3-\eta_2)\Delta_{31}^{-\theta}}{1+(\eta_2-1)\Delta_{21}^{-\theta}}\right)^{\frac{1}{\theta}} \tag{9}$$

[9]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).

Proof: see Appendix 7.1

Estimating the effect of steam power on welfare requires an estimate of the shape parameter of the Pareto distribution, θ, and an estimate of the term in large parenthesis in the rightmost expression in equation (9). With the objective of finding an expression for this term that can be estimated from firm-level data, I use the labor market clearing condition to solve for the equilibrium measure of firms, M, that are active in the economy with steam power:

$$M = \frac{\theta - \sigma + 1}{\theta \sigma f}\left(1 + (\eta_2 - 1)\Delta_{21}^{-\theta} + (\eta_3 - \eta_2)\Delta_{31}^{-\theta}\right)^{-1}$$

$$\iff 1 + (\eta_2 - 1)\Delta_{21}^{-\theta} + (\eta_3 - \eta_2)\Delta_{31}^{-\theta} = \frac{\theta - \sigma + 1}{\theta \sigma f M} \tag{10}$$

Equation (10) is recognizable as the numerator of term in the parenthesis in equation (9). Substituting this expression and imposing goods market clearing $R = wL = 1$, where R is aggregate revenue, yields:

$$\mathbb{W}/\mathbb{W}' = \left\{R/\left(\frac{\theta \sigma f M}{\theta - \sigma + 1}\left(1 + (\eta_2 - 1)\Delta_{21}^{-\theta}\right)\right)\right\}^{\frac{1}{\theta}} \tag{11}$$

The final step in deriving a welfare statistic that can be estimated from firm-level data is to find a convenient expression for the term $(\theta\sigma f M)(1 + (\eta_2 - 1)\Delta_{21}^{-\theta})/(\theta - \sigma + 1)$ in equation (11). With this objective in mind, define R_1 and R_2 as the aggregate revenue produced by firms using hand and water power respectively in the economy with steam power. Furthermore, define R_{-3} as the aggregate revenue that would have been realized by steam powered firms, had they been forced to operate with the second most productive technology (i.e. water power), holding constant the equilibrium productivity cut-offs, price index, and measure of firms at their values in the economy with steam power. Given these definitions, we have:

$$R_1 + R_2 + R_{-3} = \int_{\varphi_1}^{\varphi_2} r(\varphi, 1)\frac{g(\varphi)}{1 - G(\varphi_1)} M d\varphi + \int_{\varphi_2}^{\varphi_3} r(\varphi, \gamma_2)\frac{g(\varphi)}{1 - G(\varphi_1)} M d\varphi + \int_{\varphi_3}^{\infty} r(\varphi, \gamma_2)\frac{g(\varphi)}{1 - G(\varphi_1)} M d\varphi$$

$$= \frac{\theta\sigma f M}{\theta - \sigma + 1}(1 - \Delta_{21}^{-\theta+\sigma-1}) + \frac{\theta\sigma f M}{\theta - \sigma + 1}(\Delta_{21}^{-\theta+\sigma-1} - \Delta_{31}^{-\theta+\sigma-1})\gamma_2^{\sigma-1} + \frac{\theta\sigma f M}{\theta - \sigma + 1}\Delta_{31}^{-\theta+\sigma-1}\gamma_2^{\sigma-1}$$

$$= \frac{\theta\sigma f M}{\theta - \sigma + 1}\left(1 + (\eta_2 - 1)\Delta_{21}^{-\theta}\right)$$

Note that even though the expression above integrates over the range of firms using steam power, $\varphi \geq \varphi_3$, the definition of R_{-3} imposes that the revenue of these firms be calculated in a counterfactual environment in which they produce with water power, $\gamma = \gamma_2$. Combining this expression with equation (11) yields the following welfare equation.

$$\mathbb{W}/\mathbb{W}' = (R/(R_1 + R_2 + R_{-3}))^{\frac{1}{\theta}}\,, \tag{12}$$

In Section 4, I illustrate how this welfare equation can be estimated using firm-level data.

4 Empirical Application

4.1 Econometric Model

In this section I develop an econometric model for estimating the welfare statistic given by equation (12). The quantities R, R_1, and R_2, can be estimated by aggregate value added, and the aggregate value added produced by firms using hand and water power respectively. The empirical analogue to the theoretical quantity defined by R_{-3} is as follows:

$$\hat{R}_{-3} = \sum_{i=1}^{N_3} \hat{r}_i(\varphi_i, \gamma_2), \tag{13}$$

where N_3 is the number of firms using steam power, and $\hat{r}_i(\varphi_i, \gamma_2)$ is the revenue that would have been produced by a steam powered firm, had they been forced to use water power.

An empirical methodology for estimating $\hat{r}_i(\varphi_i, \gamma_2)$ can be derived from the theoretical model. Taking the natural logarithm of the revenue equation (4) after substituting in the monopolistically competitive price, and considering only the subset of firms using steam or water power yields:

$$ln(r(\varphi, \gamma)) = (\sigma - 1)ln(\rho P \gamma_2) + 1_{\gamma_3}(\sigma - 1)ln \quad \frac{\gamma_3}{\gamma_2} \quad + (\sigma - 1)ln(\varphi), \tag{14}$$

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:

$$y_i = \beta_0 + \beta_1 1(steam_i) + u_i, \tag{15}$$

where y_i is the natural logarithm of value added, $1(steam_i)$ is a steam power indicator variable that takes a value of one if firm i uses steam and zero otherwise, and regression equation (15) is defined over the sub-sample of firms using steam or water power. The regression coefficients and error term are defined as follows: $\beta_0 = (\sigma-1)ln(\rho P\gamma_2)$, $\beta_1 = (\sigma-1)ln(\gamma_3/\gamma_2)$, and $u_i = (\sigma-1)ln(\varphi_i)+\epsilon_i$. The error term, u_i, includes firm i's unobserved productivity type, φ_i, and the random variable ϵ_i which is assumed to be i.i.d. across firms. If a firm's productivity type is positively correlated with the adoption decision, then the coefficient β_1 will be biased upwards if estimated by OLS regression. A consistent estimate of β_1 can be obtained using instrumental variables (IV) estimation. In Section 4.2 I discuss the instruments I use to estimate equation (15) by IV estimation.

Once β_1 has been estimated, $\hat{r}_i(\varphi_i, \gamma_2)$ can be calculated for the subset of firms using steam power as follows:

$$\hat{r}_i(\varphi_i, \gamma_2) = exp(y_i - \hat{\beta}_1 1(steam_i)), \tag{16}$$

where $\hat{\beta}_1$ is the estimated value of β_1 and the remaining terms in this expression are defined in equation (15). Given an estimate of $\hat{r}_i(\varphi_i, \gamma_2)$, equation (13) can in turn be used to calculate $\hat{R}_{-3}$.

The final parameter required to estimate the welfare equation given by (12) is an estimate of the shape parameter of the Pareto distribution, θ. I now show how an estimate of θ can be obtained by estimating the parameters of the distribution of value added for the sub-sample of firms that had adopted steam power in 1871. Define $F(r)$ as the CDF of value added for the subset of firms having adopted steam power. In Appendix 7.2 I show that $F(r)$ follows a Pareto distribution with shape parameter $\theta/(\sigma-1)$. Therefore, for the subset of firms using steam power, the CDF $F(r)$ can be written:

$$ln(r) = ln(r(\varphi_3, \gamma_3)) - \frac{\sigma-1}{\theta} ln(1 - F(r)), \tag{17}$$

where $r(\varphi_3, \gamma_3)$ is the revenue of a steam powered firm with productivity equal to the steam cut-off,

φ_3. The empirical specification corresponding to equation (17) is:

$$y_i = \zeta_0 + \zeta_1 ln(1 - \hat{F}_i) + \epsilon_i \tag{18}$$

where y_i is the natural logarithm of value added, $\hat{F}_i$ is the empirical CDF, and regression equation (18) is defined over the subset of firms using steam power. I follow Head et al. (2014) in estimating equation (18) by the QQ regression methodology introduced by Kratz and Resnick (1996). Defining $i = 1$ as the firm with the minimum value added and $i = n$ as the firm with the maximum value added in the sub-sample, the empirical CDF is defined as $\hat{F}_i = (i - 0.3)/(n + 0.4)$. The coefficient of interest is $\zeta_1 = -(\sigma - 1)/\theta$. Identifying θ requires an estimate of σ, the elasticity of substitution between varieties. 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. Ziebarth (2013) assumes that $\sigma = 3$ in his analysis of misallocation and productivity of the U.S. manufacturing sector in the nineteenth century. For my analysis, I follow Melitz and Redding (2013) and Head et al. (2014) in using a benchmark parameterization of $\sigma = 4$ and explore the robustness of this assumption in the sensitivity analysis.

Once the parameter θ and the expressions R, R_1, R_2, and R_{-3} have been estimated, the welfare effect of the introduction and adoption of steam power can be calculated directly using the statistic defined by equation (12).

4.2 Identification

In theory it is straightforward to apply the estimation strategy that is presented in Section 4.1. However in practice, finding an exogenous source of variation that can be used to derive a consistent estimate of β_1 is challenging. In the case of steam power, it is possible to use exogenous variation in geography to identify the parameter β_1. In particular I use average agricultural soil quality and precipitation at the district-level as instrumental variables to estimate the effect of steam power adoption on firm-level productivity. This identification strategy draws on the theory proposed by Rothbarth (1946), Habakkuk (1962), and Clarke and Summers (1980), that agricultural land quality can explain regional differences in the pace of industrialization. In particular, these authors argue that the superior land endowment in the U.S. raised the opportunity cost of labor and encouraged the substitution of capital for labor to a greater extent in the U.S. than in the UK during the nineteenth century. In the Canadian context, McCallum (1980) has also argued that regional differences in industrial growth can be explained by agricultural land quality, although his argument relies on an alternative explanation. McCallum (1980) argues that regions of superior land quality were quicker to industrialize as a result of the capital formation financed by regional wealth accumulated in the agricultural sector during the nineteenth century. Based on the arguments of these authors, it is reasonable to hypothesize a positive correlation between firm-level steam power adoption and exogenous variation in average agricultural soil quality and precipitation levels, insofar as these measures proxy for regional agricultural endowments.

To construct a measure of average soil quality at the district-level I use geographic information systems (GIS) analysis and digital data from the Canada Land Inventory (CLI).[10] The CLI digital maps classify soil capability for agriculture using a seven-point descending scale. Class 1 is the highest agricultural soil quality classification and is described as having "no significant limitations in use for crops", (Agriculture and Agri-Food Canada, 2015). Class 7 is the lowest quality soil classification and is deemed to "have no capacity for arable culture or permanent pasture", (Agriculture and Agri-Food Canada, 2015). Using GIS analysis, I estimate average agricultural soil quality at the 1871 Canadian census district-level.[11] I also use GIS analysis to estimate the average annual

[10]The CLI was an initiative administered by the Canadian federal government's Agricultural Rehabilitation and Development Act of 1961. The ninety year lag between period of the analysis and the CLI initiative limits the accuracy of CLI data in approximating soil quality in the late nineteenth century. Despite this limitation, the CLI data provide a reasonable proxy for soil quality in 1871 as the classification was based on the potential of soils for agricultural production, rather than its observed agricultural productivity in 1961.

[11]Josh MacFadyen (University of Saskatchewan) provided the geo-spatial data for the 1871 Canadian census districts for the provinces of Ontario and Quebec. For the provinces of New Brunswick and Nova Scotia, I use the 1911 Canadian census districts, geo-edited to match the 1871 census district boundaries. The 1911 Canadian census districts are sourced from the Canadian Century Research Infrastructure project at the University of Ottawa, Canada. When calculating averages at the census-district level, I impose a two kilometer buffer around the census boundaries as a robustness measure against imprecision in the geo-referencing for the 1871 census district spatial data. While most of the census districts represented relatively large geographic areas, a number of Canadian cities were defined as self-contained census districts and consequently represented a much smaller geographic area. To

precipitation within the census districts as a second instrument to be used in IV estimation. It is important to include a measure of average annual precipitation as agricultural production depends both on soil quality and precipitation. Geo-spatial precipitation data is sourced from WorldClim Global Climate Data, which is publicly available on the internet.[12] The WorldClim data provides high geo-spatial resolution estimates of average annual precipitation over the period 1950-2000.

The instrumental variables strategy proposed above will produce a consistent estimate of β_1 if two criteria are satisfied. The first criteria requires that the instruments are highly correlated with steam power adoption. In theory, the arguments of Rothbarth (1946), Habakkuk (1962), Clarke and Summers (1980), and McCallum (1980) suggest that the instruments should be highly correlated with adoption. Empirically, this first criteria is tested by the first-stage F-statistic in two-stage least squares, which measures the correlation between the instruments and steam power adoption at the firm-level. The second criteria requires that the instruments satisfy the 'exclusion restrictions'. That is, the instruments must not be correlated with any components in the error term, u_i, that are correlated with the firm-level value added.

There are several potential violations of the exclusion restrictions that must be considered. To address these challenges I augment regression specification (15) with several control variables as follows:

$$y_i = \beta_0 + \beta_1 1(steam_i) + \mathbf{X}_i\beta + u_i, \tag{19}$$

where $\mathbf{X}_i$ is a vector of control variables. These control variables are included to address several possible challenges to the exclusion restrictions. For example, districts having the richest agricultural endowments were likely more densely populated in 1871. This higher population density may have increased firm-level output through agglomeration effects in local labor markets. To address

address this issue, I merge cities that were self-contained districts with a surrounding district to ensure greater consistency in the geographic definition of the districts.

[12]The time lag between the period of the analysis and the precipitation data limits the accuracy of WorldClim estimates in approximating precipitation patterns in the late nineteenth century. However, given that the WorldClim data were collected over a fifty-year period in the twentieth century, the data can be expected to reflect long run trends and provide a reasonable proxy for precipitation in 1871. The WorldClim data are publically available at the following url: http://www.worldclim.org/

this issue the vector of control variables includes the population density of each district as sourced from the 1871 Census of Canada. Districts with the best agricultural soils in Canada also likely had greater access to railroads in the late nineteenth century. In theory, rail access may have increased firm-level output in manufacturing through market access and external economies of scale. This issue is addressed by including a rail access indicator variable that takes a value of one if the district was serviced by a railroad. The data for the rail access variable is sourced from the 1863 Canadian railroad maps produced by Andreae (1996). Districts with high precipitation were likely to have had a higher concentration of rivers. Access to rivers may be correlated with firm-level output as many industrial activities benefited from access to running water. The vector of controls includes a variable that measures the natural logarithm of average stream flow accumulation in each district. I estimate average stream flow accumulation by GIS using a digital elevation model (DEM) with data sourced from the NASA Shuttle Radar Topography Mission (SRTM).[13] There may also have been direct benefits to manufacturing firms located in districts with rich agricultural endowments. This may have been especially true for flour and grist mills which represented an important industry in the Canadian economy during this era. This issue is partially addressed by using a value added measure of output, which nets out the value of agricultural and other intermediate inputs. However, there may have been additional benefits for firms that located in close proximity to high quality agricultural inputs that are not fully captured in the value of intermediate inputs. I address this issue by including a control variable measuring the total wheat produced per acre in each district.[14] Finally, the vector of controls also includes indicator variables for each SIC major industry group, and provincial indicator variables. These indicator variables control for industry specific technological differences and inter-provincial variation in input and output prices.

[13]SRTM DEM data was downloaded from the Consultative Group for International Agricultural Research - Consortium for Spatial Information (CGIAR-CSI) website: http://www.cgiar-csi.org/

[14]These data are sourced from the 1871 Census of Canada.

4.3 Estimation Results

The analysis in this section uses firm-level data from the 1871 Canadian Census of Manufacturing sourced from the CANIND71 database that was described in Section 2. These data are supplemented with the variables described in Section 4.2. The first step required to estimate the welfare statistic involves estimating equation (19) using the sub-sample of firms operating with either water or steam power. I estimate equation (19) by OLS, two variants of instrumental variables estimation, and the control function methodology proposed by Wooldridge (2015).

The first IV estimator used is the standard two-stage least squares methodology (henceforth, 2SLS). The second IV estimator is the *Procedure 18.1* approach proposed by Wooldridge (2002). The first step in Wooldridge's *Procedure 18.1* estimator (henceforth, IV-WP) involves regressing the steam power indicator variable on the full set of instruments and the control variables from the equation (19) using a probit regression.[15] The fitted probabilities are then obtained from the probit regression results and are used as an instrumental variable in estimating equation (19) by two-stage least squares. The IV-WP estimator is more efficient than standard two-stage least squares estimation, however this result relies on the probit regression model being correctly specified in the first step, (Wooldridge, 2002).[16] Importantly, the IV-WP estimator remains consistent and the standard errors of the coefficient estimates are asymptotically valid, even when probit regression is misspecified.

The control function estimator (henceforth, CF) entails running a first-stage probit regression of the steam power indicator variable on the full set of instruments and the control variables from the equation (19). The fitted values from the first stage probit are then used to calculate the 'generalized residual'.[17] A consistent estimate of β_1 can then be obtained by running a second

[15]More generally, the first step in Wooldridge's *Procedure 18.1* can be estimated by any binary response model using maximum likelihood.

[16]In the context of the theoretical model developed in this paper, a probit specification of the binary response model would be correctly specified if I assumed that ex-ante firm-level productivity followed the log-normal distribution instead of the Pareto distribution. Head et al. (2014) argue that there are theoretical and empirical justifications for using the log-normal distribution as an alternative to the Pareto distribution when modeling firm-level productivity.

[17]The generalized residual has the following functional form $r(1(steam_i), \mathbf{W_i}\delta) = 1(steam_i)\lambda(\mathbf{W_i}\delta) - (1 -$

stage regression that is specified by equation (19), augmented with the inclusion of the generalized residual as an additional control variable. However, the consistency of the CF estimator depends on the probit regression being correctly specified. In this respect, the CF estimator is less robust than the IV-WP and 2SLS estimators. Bootstrap standard errors for the CF estimates are calculated using one thousand bootstrap replication samples.

Table 3: Regression of the Natural Logarithm of Value Added on Steam Power Adoption

Variable/Estimate	OLS	2SLS	IV-WP	CF
Steam Power	1.123***	0.600***	0.423***	0.951***
	(0.0408)	(0.168)	(0.142)	(0.0423)
Population Density	1.636***	1.785***	1.836***	1.585***
	(0.0757)	(0.0897)	(0.0856)	(0.0735)
Rail Access	0.0638**	0.0896***	0.0983***	0.0424
	(0.0324)	(0.0332)	(0.0332)	(0.0312)
ln(Stream Flow)	0.172***	0.176***	0.178***	0.163***
	(0.0156)	(0.0158)	(0.0159)	(0.0154)
Wheat/Acre	0.271***	0.306***	0.318***	0.252***
	(0.0493)	(0.0497)	(0.0495)	(0.0499)
γ_3/γ_2	1.454***	1.221***	1.151***	1.373***
	(0.0198)	(0.0686)	(0.0545)	(0.0189)
$R/(R_1 + R_2 + R_{-3})$	1.393	1.233	1.168	1.345
Observations	9,009	9,009	9,009	9,009
R-squared	0.418	0.406	0.397	0.427
Kleibergen-Paap rk Wald F-statistic		231.4		

*** p<0.01, ** p<0.05, * p<0.1. For γ_3/γ_2 the asterisks represent statistical significance for a test of the null hypothesis that this ratio is equal to one. Robust standard errors in parentheses. For the control function estimates (CF), bootstrapped standard errors calculated from 1000 bootstrap replications are reported. Indicator variables for SIC major industry groups and province of establishment are included in all specifications. Coefficient estimates for these variables are not reported but are available upon request. The estimate of γ_3/γ_2 is calculated using a parameterized value of $\sigma = 4$.

Table 3 reports the OLS, 2SLS, IV-WP, and CF regression results along with the estimates

$1(steam_i))(\lambda(-\mathbf{W_i}\delta))$, where the term $\mathbf{W_i}\delta$ is the fitted values from the first stage probit regression, and λ is the inverse Mills ratio, $\lambda = \phi(.)/\Phi(.)$ with $\phi(.)$ and $\Phi(.)$ being the pdf and CDF of the standard normal distribution.

of γ_3/γ_2 and the ratio $R/(R_1 + R_2 + R_{-3})$. The estimates of γ_3/γ_2 use the parameterized value of $\sigma = 4$, and are derived from the coefficient estimate for the steam power indicator variable, $\beta_1 = (\sigma - 1)ln(\gamma_3/\gamma_2)$. The standard errors for the ratio γ_3/γ_2 are calculated using the delta method. To estimate the ratio $R/(R_1 + R_2 + R_{-3})$, the terms R, R_1, and R_2 are calculated from the census data on aggregate value added, and the value added produced by firms using hand and water power respectively. The term R_{-3} is calculated from the census data on steam powered firms combined with the regression results, and using equations (13) and (16).

The results in Table 3 suggest a positive and statistically significant relationship between steam power adoption and value added. The coefficient estimates for the steam power indicator variable are positive and statistically significant at the 1 percent level in every column of Table 3. This is not surprising given the results that were reported in Section 2. More interesting is the decline in the estimates of β_1 and γ_3/γ_2 between OLS and the IV and CF estimators. Recall that the ratio γ_3/γ_2 measures the gross increase in firm-level productivity from using steam power relative to water power. The 2SLS, IV-WP, and CF estimates of this parameter are 23, 30, and 8 percentage points lower than the OLS estimate respectively. This is consistent with the notion 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 22.1, 15.1, and 37.3 percent increase in firm-level labor productivity for the 2SLS, IV-WP, and CF estimates respectively.[18] 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 four times that of water powered firms.

The estimates for the ratio $R/(R_1 + R_2 + R_{-3})$ are also reported in Table 3. The 2SLS, IV-WP, and CF derived estimates of this ratio show a marked declined from the OLS estimates. I choose to use the IV-WP estimate of $R/(R_1 + R_2 + R_{-3})$ in calculating the welfare statistic as it is the most conservative of the estimates.

For the 2SLS estimates, the Kleibergen-Paap rk Wald F-statistic is reported at the bottom of Table 3. The null hypothesis of weak-identification is soundly rejected at any reasonable level of

[18]These estimates are statistically significant at the 1 percent level.

significance. In Table 7 in Appendix 7.3 I also report the first-stage regression results for the 2SLS estimator, and the probit regression used in the IV-WP and CF estimators. Both soil quality and precipitation are highly significant and of the expected sign, providing further evidence in support of the strength of the identification strategy. In particular, the results suggest a strong positive correlation between firm-level adoption of steam power and regional agricultural endowments, as measured by soil quality and precipitation.

Table 4: QQ Regression Results and Estimates of the Welfare Effect of Steam Power Adoption

Estimate	All	Top 50	Top 25	Top 20	Top 15	Top 10	Top 5
$-(\sigma-1)/\theta$	-1.362	-0.977	-0.811	-0.766	-0.717	-0.658	-0.574
	(0.0110)	(0.00546)	(0.00607)	(0.00629)	(0.00627)	(0.00641)	(0.00803)
θ	2.202	3.069	3.701	3.914	4.184	4.562	5.229
	(0.0177)	(0.0171)	(0.0277)	(0.0321)	(0.0366)	(0.0445)	(0.0732)
$\mathbb{W}/\mathbb{W}'$		1.052	1.043	1.041	1.038	1.035	1.030
Observations	2,630	1,349	658	526	408	282	133
R-squared	0.855	0.960	0.964	0.966	0.970	0.974	0.975

Standard errors are in parentheses. The estimates of the shape parameter, θ, and the welfare statistic, $\mathbb{W}/\mathbb{W}'$, are calculated using $\sigma = 4$, and the IV-WP derived estimate of $R/(R_1 + R_2 + R_{-3})$ from Table 3.

The next step required to estimate the welfare statistic is to estimate the shape parameter of the Pareto distribution, θ. The QQ regression estimation results of equation (18) are reported in Table 4. I follow Head et al. (2014) in reporting results for the entire sample and several sub-samples that are created by selecting the right tail of the empirical distribution at various percentiles. Head et al. (2014) note that the Pareto distribution is commonly used in the empirical trade literature because it is tractable, can be theoretically micro-founded, and empirically provides a good fit of the right tail of the observed distribution of manufacturing firm sales data. However, the left tail of this distribution is not as well approximated by the Pareto distribution, and it is not uncommon for researchers to estimate values of the shape parameter that violate the key assumption that was noted in Section 3, that $\theta > \sigma - 1$. Table 4 illustrates that this assumption is indeed violated for the full sample. Theoretically valid estimates of the shape parameter, θ, are obtained when the sample is restricted to the top 50 percent or less (as ordered by size in value added). I follow Head et al. (2014) in reporting the welfare statistic for all theoretically valid estimates of θ.

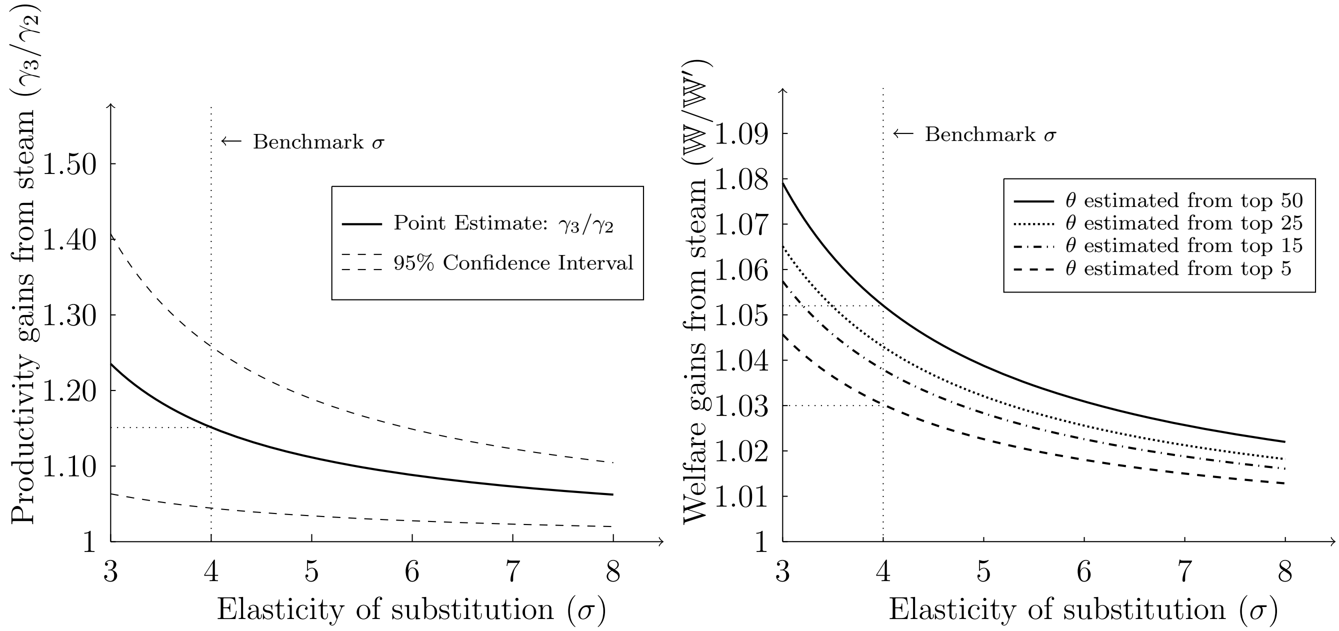

Figure 3: Firm-Level Productivity and Aggregate Welfare Gains from Steam Power Adoption

In Table 4 the estimates of the shape parameter, θ, use the parameterized value of $\sigma = 4$, and the estimated value of the parameter $-(\sigma - 1)/\theta$ from the QQ regression. The welfare statistic, $\mathbb{W}/\mathbb{W}'$, is then estimated using the equation (12) and the estimates of θ and the IV-WP estimate of the ratio $R/(R_1 + R_2 + R_{-3})$. The results of this calculation for the various estimates of θ are presented in Table 4. The results indicate that the introduction and adoption of steam power resulted in an increase in welfare that ranged from 3.0-5.2 percent depending on the estimate of θ that is used.

As previously noted, the framework developed in this paper 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. Melitz and Redding (2013) use values of $\theta = 4.25$ and $\sigma = 4$ in their calibration. In Table 4, the estimated value of θ for the top 15% of firms is 4.184, which is close to the value used by Melitz and Redding (2013). Using this estimate of θ yields an estimated aggregate welfare gain of 3.8 percent from steam power adoption. Thus, the estimated welfare gains from steam power adoption are roughly one and a half times 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 parameterized value of the elasticity of substitution, σ. Figure 3 explores 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 left panel of Figure 3 plots the point estimate and 95% confidence interval for the increase in firm-level productivity from firms switching from water to steam power. The estimated productivity gains are decreasing in σ and range from 23.5 to 6.2 percent. The right panel plots the estimated aggregate welfare gain from the introduction and adoption of steam power in the Canadian manufacturing sector in 1871. The various line plots correspond to the different sub-samples used in estimating the shape parameter θ. The estimated welfare gain from steam power adoption is also decreasing in σ and ranges from a high of 7.9 percent to a low of 1.3 percent.

5 Firm Heterogeneity in GPT Adoption: the Case of Internet-Enabled Mobile Devices

The results in Section 4 suggest that the welfare gains from steam power were roughly one and a half times as large as contemporary estimates of the welfare gains from trade. Given the methodological similarities, comparing the welfare gains from steam power adoption to the gains from trade is a useful means of benchmarking the economic significance of these estimates. However, quantifying the welfare gains from steam relative to other GPTs requires analysis using firm-level data on the adoption of different GPTs from different eras. To achieve this objective, I use enterprise-level data on internet-enabled mobile device adoption (i.e. smartphones and tablets) in the Canadian private sector in the twenty-first century. This section presents a number of empirical results related to the use of mobile devices that demonstrate that the pattern of adoption is consistent with the theoretical framework developed in Section 3.

Table 5: Internet-Enabled Mobile Device Adoption Rate by Size and Sector in Canada, 2012

Size\Sector	11	21	22	23	31-33	41	44-45	48-49	51	52
Small	30.1			47.6	52.4	64.2	42.9	39	74.9	47.2
Medium	47.7			83.2	78.1	87.6	62.7	72.6	79.1	89.8
Large	84.4	100	100	90.1	98.4	91.6	71	93.5	97.8	99.7
All	31.4	64.4	77.2	50.3	60.1	67.4	45.8	41.2	76.7	53.2

Size\Sector	53	54	55	56	61	62	71	72	81	Private Sector
Small	48.9	68.2	39	52.8	53.5	46.3	59.2	38.6	38.4	50.3
Medium	81.7	93.3	74.2	79.2	77.3	71.1	70.7	54.6	52.4	74.4
Large	99.9	94.6	100	79.8	94.5		94.2	87.5	99.7	88
All	51.5	69.7	52.9	55.2	56.4	48.2	60.4	41.4	39.9	53

Source: Statistics Canada CANSIM Table: 358-0197. NAICS two-digit industry codes are defined as follows: 11: agriculture, forestry, fishing and hunting; 21: mining, quarrying, and oil and gas extraction; 22: utilities; 23: construction; 31-33: manufacturing; 41: wholesale trade; 44-45: retail trade; 48-49: transportation and warehousing; 51: information and cultural industries; 52: finance and insurance; 53: real estate and rental and leasing; 54: professional, scientific and technical services; 55: management of companies and enterprises; 56: administrative and support, waste management and remediation services; 61: educational services; 62: health care and social assistance; 71: arts, entertainment and recreation; 72: accommodation and food services; 81: other services (except public administration). The Private Sector includes all NAICS two-digit industries 11-81. For all sectors except manufacturing, the enterprise size categories are defined as follows: Small: 0-19 employees; Medium: 20-99 employees; Large: 100 or more employees. For manufacturing, the enterprise size categories are defined as follows: Small: 0-19 employees; Medium: 20-499 employees; Large: 500 or more employees. Missing values have been suppressed by Statistics Canada to meet the confidentiality requirements, or because the estimates are too unreliable to be published.

The enterprise-level data on internet-enabled mobile device adoption comes from Statistics Canada's Survey of Digital Technology and Internet Use (SDTIU). Table 5 presents the average adoption rate of mobile devices for each sector, cross-tabulated by enterprise-size in the year 2012.[19] The data show a positive correlation between the use of mobile devices and enterprise-scale that is ubiquitous across the Canadian private sector. In every sector there is a monotonic relationship between the three size categories and the adoption rate. The overall adoption rate in the Canadian private sector is 53%, although nearly all (88%) of large enterprises use internet-enabled mobile devices. This positive correlation between scale and adoption is consistent with the pattern for steam power, as highlighted in Table 1 and Figure 1. This correlation is also consistent with the theoretical framework, which assumes that firm-level heterogeneity and the fixed-costs of adoption lead only the largest and most productive firms to use frontier technologies.

To analyze the relationship between productivity and mobile adoption, I link the SDTIU to administrative data that records sales, inputs, and an array of other enterprise-level characteristics. Table 6 presents results from OLS regressions of labor productivity on mobile device adoption. For the regression analysis I pool the SDTIU and administrative data over the years 2012-2013.[20] Regressions are run for each major sector of the economy: manufacturing, services, and other.[21] Columns (7) and (8) in Table 6, combine all enterprises in the sample and thus provide estimates for the private sector. The dependent variable is the natural logarithm of labor productivity, measured

[19]For all sectors except manufacturing, the enterprise size categories are defined as follows: Small: 0-19 employees; Medium: 20-99 employees; Large: 100 or more employees. For manufacturing, the enterprise size categories are defined as follows: Small: 0-19 employees; Medium: 20-499 employees; Large: 500 or more employees.

[20]In each regression I cluster the standard errors at the enterprise level. The cross-sectional survey weights from 2012 are applied to both the 2012 and 2013 sample. In particular, enterprises surveyed in both 2012 and 2013 are assigned their 2012 survey weight for both observations. Enterprises surveyed only in 2012 are assigned their 2012 survey weight. Enterprises surveyed only in 2013 are excluded from the sample. As a robustness check, all regressions were also estimated without using sampling weights. The qualitative results are unaffected by the use of the sampling weights while the quantitative results vary only slightly.

[21]Manufacturing includes NAICS two-digit industries 31-33, services includes NAICS two-digit industries 41-81, and other includes NAICS two-digit industries 11-23. See Table 5 for definitions of each NAICS two-digit industry.

Table 6: Regression Analysis of Labor Productivity for Canadian Private Sector, 2012-2013

	Manufacturing		Services		Other		Private Sector	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	ln(VA/L)	ln(PQ/L)	ln(VA/L)	ln(PQ/L)	ln(VA/L)	ln(PQ/L)	ln(VA/L)	ln(PQ/L)
Mobile Devices	0.0914**	0.154***	0.144***	0.110**	-0.208*	-0.120	0.116***	0.101**
	(0.0410)	(0.0460)	(0.0394)	(0.0475)	(0.115)	(0.154)	(0.0351)	(0.0429)
Fewer than 20 Employees	-0.102***	-0.112**	0.0920***	0.0756*	-0.222**	-0.387**	0.0574*	0.0208
	(0.0370)	(0.0466)	(0.0339)	(0.0452)	(0.0971)	(0.161)	(0.0295)	(0.0383)
ln(District Population)	0.0435***	0.0879***	0.0260	0.0296	0.102**	0.106*	0.0297**	0.0372*
	(0.0156)	(0.0188)	(0.0165)	(0.0253)	(0.0435)	(0.0558)	(0.0146)	(0.0221)
ln(Capital/Labor)	0.120***	0.151***	0.108***	0.0673***	0.166***	0.198***	0.114***	0.0820***
	(0.0165)	(0.0224)	(0.0146)	(0.0168)	(0.0322)	(0.0469)	(0.0133)	(0.0152)
Province Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NAICS 3 Indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2012 Indicator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,528	2,719	10,198	9,775	795	872	13,521	13,366
R-squared	0.264	0.185	0.414	0.289	0.330	0.188	0.397	0.276

Standard errors in parentheses have been clustered at the enterprise level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Manufacturing includes NAICS two-digit industries 31-33, Services includes NAICS two-digit industries 41-81, and Other includes NAICS two-digit industries 11-23. The Private Sector includes all NAICS two-digit industries 11-81. See Table 5 for definitions of each NAICS two-digit industry. Coefficient estimates for province, NAICS three-digit indicator, and year 2012 indicator variables are not reported but are available upon request.

either by value-added divided by labor input, or by sales divided by labor input.[22] A measure of intermediate inputs is not available and therefore I approximate value-added by the sum of net income, wages and salaries, and amortization expenses.[23] The second specification of labor productivity (sales divided by labor input) is included to show the robustness of the results to any measurement error associated with the approximation of value added. The key variable of interest in the regression analysis is the mobile device indicator variable. This variable takes a value of one if the enterprise indicated that it used internet-enabled mobile devices, and zero otherwise. The right-hand side variables include the natural logarithm of the capital-labor ratio, which controls for capital intensity.[24] Each regression in Table 6 also includes controls for small enterprises (i.e. those having fewer than twenty employees), the population in the enterprise's census district, indicator variables for the enterprise's province of residence, indicator variables for each enterprise's NAICS three-digit industry, and an indicator variable that takes a value of one if the year of the observation is 2012.

The results in Table 6 indicate that internet-enabled mobile device adoption is correlated with labor productivity in the manufacturing and service sectors of the Canadian economy. The coefficient on the mobile device indicator is positive and statistically significant at the five percent level in each specification for both of these sectors. The coefficient is negative for the other industries sector, however it is not statistically significant at the five percent level in either specification. When the regressions are run for the entire private sector the coefficient of interest is positive and statistically significant at the five percent level in each specification.

It is interesting to compare the regression results from Table 2 with those for the manufacturing sector in Table 6. In particular, column (3) of Table 2 provides a comparable specification to column

[22] Enterprise-level sales are sourced from Statistics Canada's General Index of Financial Information (GIFI) database. Labor input is measured by individual labor units (ILUs), which are sourced from Statistics Canada's Longitudinal Employment Analysis Program (LEAP) database.

[23] This specification of value added follows from the definition of net income: *net income* = *value added* − *wages and salaries* − *amortization expenses*. Net income and amortization expenses are sourced from the GIFI database, and wages and salaries are sourced from the LEAP database.

[24] Capital is measured by the enterprise's tangible capital assets, net of accumulated amortization. These variables are sourced from the GIFI database.

(1) in Table 6. For the manufacturing sector, the magnitude of the coefficient estimates for steam power in the nineteenth century is comparable to that of mobile devices in the twenty-first century. The magnitude, sign and significance of the small firm, district population, and capital-labor ratio are also comparable across column (3) of Table 2 and column (1) of Table 6. These similarities are suggestive of long-run persistence in internal and external scale economies in the Canadian manufacturing sector.

In summary, the results in this section show a similarity between the characteristics of enterprises using mobile devices in the twenty-first century and the characteristics of firms using steam power in the late nineteenth century. In both cases adoption is much more common amongst large firms, and the pattern of adoption indicates a positive correlation with labor productivity. The theoretical model developed in Section 3 is consistent with these empirical results, and it provides a framework for quantifying the relative productivity and welfare gains associated with steam and internet-enabled mobile devices. The objective of my current research is to use this framework to estimate the productivity and welfare gains from mobile devices, and to compare the relative gains from these two GPTs.

6 Conclusion

In this paper I develop a tractable framework for calculating the firm-level productivity and aggregate welfare gains that are accrued from the introduction and adoption of a new technology. My analysis of steam power adoption in the late nineteenth century illustrates the potential of my model, and at the same time contributes to the literature on general purpose technologies. My results suggest that the introduction of a GPT can have significant impacts on aggregate welfare, even during the early stages of diffusion when only a small fraction of the economy has adopted the technology. Despite the fact that only 8 percent of Canadian manufacturing firms used steam in 1871, my results suggest that the adoption of steam power resulted in a 3.0-5.2 percent increase in welfare. This finding can be partially explained by the fact that firms using steam were large, having accounted for approximately 42 percent of Canadian manufacturing value added in 1870. A second factor explaining the large welfare gains is the estimated increase of 15.1 percent in firm-level productivity for firms using steam power relative to those using water power. Therefore, this paper finds that technological adoption by a relatively small number of firms can have a pronounced effect on aggregate welfare when positive selection results in a reallocation of resources to the highly productive firms that are the first to make use of the technology.

The theoretical framework developed in this paper has a parsimonious structure, yet is able to replicate the empirical regularities of steam power adoption in nineteenth century. The methodology can also be extended to study other research questions pertaining to firm heterogeneity in technological adoption. For example, Ristuccia and Solomou (2014) find that the productivity benefits of GPT adoption vary across industries. The theoretical framework in this paper could be easily extended to estimate the sector-specific welfare gains from GPT adoption, in a similar fashion to the approach used in multi-sector models of international trade, such as Melitz and Redding (2014). The methodology in this paper can also be applied to study the welfare gains from other GPTs such as electricity and ICTs. Crafts (2004) finds that relative to ICTs, the contribution of steam technology to economic growth was modest. However the growth accounting framework used in Crafts (2004) ignores the productivity and welfare effects associated with firm heterogeneity. When these effects are incorporated in the theoretical framework, this paper finds that the welfare gains from steam power were roughly one and a half times as large as contemporary estimates of the welfare gains from trade. My current research is focused on using the framework of this paper to measure the productivity and welfare gains from internet-enabled mobile devices. The preliminary results presented in Section 5 show qualitative similarities in the pattern of adoption for steam power and mobile devices at the firm-level. My future research on this topic will extend this comparison by using the theoretical framework of this paper to quantify the productivity and welfare gains from mobile devices relative to steam power.

References

Agriculture and Agri-Food Canada (2015) 'Overview of classification methodology for determining land capability for agriculture.' *Canada Land Inventory, http://sis.agr.gc.ca/cansis/nsdb/cli/classdesc.html*

Andreae, Christopher (1996) *Lines of Country: An Atlas of Railway and Waterway History in Canada* (Boston Mills Press)

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

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

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

Clarke, Richard N., and Lawrence H. Summers (1980) 'The labour scarcity controversy reconsidered.' *Economic Journal* 90(357), 129–139

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

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

Habakkuk, Hrothgar John (1962) *American and British technology in the nineteenth century* (Cambridge University Press)

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

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

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

McCallum, John (1980) *Unequal Beginnings. Agriculture and Economic Development in Quebec and Ontario until 1870.* (University of Toronto Press)

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

Ristuccia, Cristiano Andrea, and Solomos Solomou (2014) 'Can general purpose technology theory explain economic growth? electrical power as a case study.' *European Review of Economic History* 18(3), 227–247

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

Rothbarth, Erwin (1946) 'Causes of the superior efficiency of USA industry as compared with British industry.' *Economic Journal* 56(223), 383–390

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

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

_ (2015) 'Control function methods in applied econometrics.' *The Journal of Human Resources* 50(2), 420–445

Ziebarth, Nicolas L. (2013) 'Are China and India backward? Evidence from the 19th century U.S. Census of Manufactures.' *Review of Economic Dynamics* 16, 86–99

7 Appendix

7.1 Proof of Proposition 1

It is convenient to invert the demand curve and write the revenue of a firm operating with hand power as a function of demand:

$$r(\varphi, 1) = P^{\frac{\sigma-1}{\sigma}} q(\varphi, 1)^{\frac{\sigma-1}{\sigma}} \tag{20}$$

The above expression also defines firm-level revenue in the economy without steam power. Note that the revenue of a firm operating with productivity equal to the operating cut-off is equal in the economy with and without steam, $r(\varphi_1, 1) = r(\varphi_1^{'}) = \sigma f$. Using this equality to evaluate equation (20) at φ_1 and $\varphi_1^{'}$ yields:

$$\mathbb{W}/\mathbb{W}^{'} = P^{'}/P = q(\varphi_1, 1)/q(\varphi_1^{'}, 1) \tag{21}$$

Finally, I use the fact that $r(\varphi_1, 1) = r(\varphi_1^{'}) = \sigma f$ along with the definition of CES demand implies $q(\varphi_1, 1) = (\sigma - 1) f \varphi_1$ and $q(\varphi_1^{'}, 1) = (\sigma - 1) f \varphi_1^{'}$. Combining these expressions with equation (21) and substituting in the equilibrium values of the operating cut-offs, φ_1 and $\varphi_1^{'}$, completes the proof. $\square$

7.2 Derivation of the Distribution of Value Added for Firms with Steam Power

Define $F(r)$ as the CDF of value added for the subset of firms using steam power. By definition:

$$F\left(r\right) = Pr\left(r(\varphi, \gamma_3) \leq r \mid r(\varphi_3, \gamma_3) \leq r(\varphi, \gamma_3)\right) = \frac{Pr\left(r(\varphi_3, \gamma_3) \leq r(\varphi, \gamma_3) \leq r\right)}{Pr\left(r(\varphi, \gamma_3) \geq r(\varphi_3, \gamma_3)\right)},$$

$$= \frac{Pr\left(r(\varphi, \gamma_3) \leq r\right) - Pr\left(r(\varphi, \gamma_3) \leq r(\varphi_3, \gamma_3)\right)}{Pr\left(r(\varphi, \gamma_3) \geq r(\varphi_3, \gamma_3)\right)} \tag{22}$$

Evaluating the revenue equation (4) at the monopolistically competitive price yields: $r(\varphi, \gamma_3) = (\rho P \gamma_3)^{\sigma-1} \varphi^{\sigma-1}$. Using this expression and equation (22), the distribution $F(r)$ for $r \geq r(\varphi_3, \gamma_3)$ is:

$$F\left(r\right) = \frac{Pr\ \ (\rho P \gamma_3)^{\sigma-1} \varphi^{\sigma-1} \leq r\ \ - Pr\ \ (\rho P \gamma_3)^{\sigma-1} \varphi^{\sigma-1} \leq (\rho P \gamma_3)^{\sigma-1} (\varphi_3)^{\sigma-1}}{Pr\ \ (\rho P \gamma_3)^{\sigma-1} \varphi^{\sigma-1} \geq (\rho P \gamma_3)^{\sigma-1} (\varphi_3)^{\sigma-1}},$$

$$= \frac{Pr\left(\varphi \leq \frac{r^{\frac{1}{\sigma-1}}}{\rho P \gamma_3}\right) - Pr\left(\varphi \leq \varphi_3\right)}{Pr\left(\varphi \geq \varphi_3\right)} = \int_{\varphi_3}^{\frac{r^{\frac{1}{\sigma-1}}}{\rho P \gamma_3}} \frac{g(\varphi)}{1 - G(\varphi_3)} d\varphi \tag{23}$$

It follows that the CDF of value added for the subset of firms that have adopted steam power is:

$$F(r) = \begin{cases} 0 & \text{if } r < r(\varphi_3, \gamma_3) \\ 1 - (r(\varphi_3, \gamma_3))^{\frac{\theta}{\sigma-1}} r^{-\frac{\theta}{\sigma-1}} & \text{if } r \geq r(\varphi_3, \gamma_3) \end{cases} \tag{24}$$

7.3 Auxiliary Regressions used in Instrumental Variables Estimation

Table 7: Regression of Steam Power Adoption on Instruments and Control Variables

Variables	OLS	Probit
Agricultural Soil Quality	-0.0810***	-0.306***
	(0.00438)	(0.0186)
Precipitation	0.000792***	0.00328***
	(6.13e-05)	(0.000309)
Population Density	0.193***	0.747***
	(0.0241)	(0.0994)
Rail Access	0.0400***	0.230***
	(0.00863)	(0.0395)
ln(Stream Flow)	0.0151***	0.0641***
	(0.00415)	(0.0161)
Wheat/Acre	-0.0633***	-0.264***
	(0.0185)	(0.0615)
Province Controls	Yes	Yes
SIC Industry Controls	Yes	Yes
Observations	9,009	9,009
R-squared	0.361	

Robust standard errors in parentheses. Coefficient estimates for province and industry indicator variables are not reported but are available upon request.