A direct test of the Stolper-Samuelson theorem: the natural experiment of Japan[1]

Daniel M. Bernhofen

University of Nottingham, CESifo and GEP

John C. Brown

Clark University and GEP

Tanimoto Masuyuki

University of Tokyo

March 5, 2013

Abstract:

We exploit a natural experiment to test a classic proposition in economics: the Stolper-Samuelson theorem. We assert that Japan's 19th century trade liberalization allows us to observe the responses of an economy's factor prices to changes in goods prices in a time period long enough permitting factors to reallocate in response to these goods price changes but also short enough to fulfill the critical ceteris paribus assumptions of the theory. Employing a unique historical set which matches market level goods and factor price data with the corresponding matrix of direct and indirect factor input requirements, we fail to reject Stolper-Samuelson.

Key words: testing Stolper-Samuelson; natural experiment of a Walrasian economy.

JEL classification: F11, F14, N10, N75.

---

[1] Addresses for Correspondence: Daniel Bernhofen, School of Economics, University of Nottingham, University Park, Nottingham, NG7 2 RD, UK. Tel: 44 115 846 7055, Fax: 44 115 951 4159, email: daniel.bernhofen@nottingham.ac.uk. Earlier versions of this paper were presented at a 2011 GEP-CESifo joint conference in Munich, a 2012 Madrid conference on 'Trade, poverty and growth in history' and at department seminars at Clark, Columbia, the University of Tokyo and Yale University. We acknowledge the generous support of the National Science Foundation Grant SES-0452991 and the Leverhulme Trust Grant F/00 114/AM. Shōji Masahiro provided invaluable research assistance. Gbeton Somasse contributed significantly to the formal statistical tests.

1. **Introduction**

Since it was first formulated over 70 years ago, the Stolper and Samuelson (1941) theorem has remained one of the most fundamental propositions of general equilibrium (trade) theory. The theorem predicts how trade-induced changes in goods prices affect changes in factor prices in a Walrasian competitive equilibrium under the assumptions of unchanged technologies, factor supplies and preferences. Although Stolper-Samuelson has served as the key framework for examining the general equilibrium effects of trade liberalization on domestic factor prices, the implausibility of the theorem's critical assumptions regarding modern globalization experiences has constrained efforts to conduct a refutable test of it.

This paper breaks new ground by exploiting Japan's 19th century move from autarky to open trade as a natural experiment that permits a direct test of the higher dimensional formulation of the Stolper-Samuelson prediction. The forced opening up of Japan in 1859 provides an unusual opportunity to observe the impact of trade induced changes in commodity prices on factor prices in what can be categorized as a textbook type Walrasian island economy. The economy's initially slow adaption of western technologies have allowed us to identify experimental 'autarky' and 'open trade' windows separated by 20-25 years. We argue that the time period between these two experimental windows was long enough permitting factors to reallocate in response to the exogenous goods price changes but also short enough to fulfil the critical ceteris paribus assumptions of the theory.

We test the most general formulation of the theorem, which was developed by Ethier (1982) and Ethier (1984), to this experimental setting. Ethier generalized Stolper and Samuelson's (1941) original two-good, two-factor formulation to any number of goods and factors. Ethier's formulation offers a refutable prediction on the responses of factor prices to (trade-induced) changes in goods prices: a factor intensity weighted change in the economy's goods price vector imposes a restriction on the change in the economy's factor price vector. This higher dimensional formulation is a *system prediction* on the average changes of *all factor prices* in the economy. It says that commodity price changes, on average, raise the most the rewards of the factors employed most intensively in the goods for which the prices have changed the most. A virtue of this formulation is that it is a falsifiable general equilibrium prediction that is not conditional on dimensionality or on specific functional forms for technologies or consumer

preferences. Since this formulation accommodates environments in which factors are specific to individual goods or subset of goods, it encompasses also the specific factor model.

In a Walrasian economy, the direction of factor price movements is central to the working of the invisible hand in reallocating factors following a change in goods prices. Assume for example, that an exogenous shock rises the relative price of good *j*, ceteris paribus. The resulting expansion of good *j* requires a corresponding movement of factors to that sector. But which factors should move to guarantee the maximum output of all goods in the economy? Stolper-Samuelson postulates that the factors that are used most intensively in good j should experience the largest price increases to entice them to move. So commodity price changes interacted with relative factor intensities impose a restriction on the permissible changes in factor prices that ensure an efficient reallocation of resources. An empirical confirmation of this formulation of Stolper-Samuelson provides evidence on the working of decentralized goods and factor markets in solving the economy's resource allocation problem.[2]

Our paper is part of an emerging literature exploiting natural experiments in economic history to seek empirical evidence on causal mechanisms regarding the behavior of the economic system. Since the corresponding systems predictions pertain to complex social phenomenon, as described in the classic treatment by von Hayek (1964), empirical verification is beyond the scope of lab, field or random experiments. An attractive feature of this new literature is that the underlying theoretical framework

[2] Our paper relates also to the literature on the testability of general equilibrium theory. In two seminal papers, Sonnenschein (1973) Sonnenschein (1974) asked whether general equilibrium theory could generate non-trivial testable restrictions on aggregate excess demand for an exchange economy. Follow-up work by Mantel (1974)and Debreu (1974)established the so-called Sonnenschein-Mantel-Debreu (SMD) theorem which asserts that the response of aggregate excess demand to price changes can be arbitrary. The SMD or "anything can happen" theorem has been interpreted to imply that the Walras-Arrow-Debreu general equilibrium model does yield testable restrictions. Brown and Matzkin (1996)) have challenged the SMD perspective by pointing out that aggregate excess demand might not be the appropriate object on which to impose testable restrictions. Specifically, Brown and Matzkin launched a line of theoretical inquiry deriving restrictions on the *equilibrium manifold*, defined as the set of prices and endowments at which excess demand is zero. It seems to have gone unnoticed that Ethier's (1982 and 1984) general Stolper-Samuelson formulation yields testable restriction on equilibrium factor price responses that hold under very general conditions. By testing this prediction we provide evidence in favour of the empirical validity of Walrasian general equilibrium theory per se.

guarantees *internal* validity while the natural experimental implementation guarantees a relatively high degree of *external* validity of the empirical findings. Although the experiment is conducted within a specific historical event or time period, the focus is on the testing of a specific mechanism which is not specific to that historical event or period. Recent examples include Redding and Sturm (2008)'s exploitation of German division after World War II to test systems predictions on market access, Ahlfeld, Redding et al. (2012)'s exploitation of the building and dismantling of the Berlin Wall to test system theories on location choice and Donaldson (2013)'s study of railroad construction in colonial India to examine patterns of regional specialization of welfare.

Our empirical analysis exploits a unique historical data set that is compatible with the theoretical requirement of observing market level price changes of the same (18 in our case) homogenous goods. Prices are drawn from detailed sources of monthly transaction prices spanning the time period of 1850 to 1879. Our historical data set has been constructed to identify all direct and indirect unit factor inputs required for each good. Our research on the nature of Japanese factor markets at the time has resulted in a 5-factor classification: skilled male workers, unskilled male workers, female workers, capital and land. The corresponding factor prices are drawn from historical sources and build on the extensive groundwork laid by Japanese economic historians.

A major innovative component of the data analysis is the employment of a direct and indirect factor input requirement matrix for the 18 products and 5 factors of production. A unique and highly attractive feature of this technology matrix is that the economy-wide input-output relationships are measured at the product level, rather than pertaining to industry aggregates which characterize 20^{th} century economies. Our product-level technology matrix is constructed from direct input requirements at the task level and also incorporates factor requirements of the key intermediate inputs. It draws on a range of historical sources, including a major Japanese survey of agricultural techniques, accounts by European observers and numerous studies by Japanese and western scholars that draw on village records and business accounts.

The nature of the theoretical formulation is flexible enough to permit us to test the prediction under different assumptions regarding the regional segmentation of the

market for land. Our empirical implementation also consider two alternative experimental free trade windows (1871-75 and 1876-79) to ensure that our measures of price changes reflect changes in 'long-term' fundamentals rather than cyclical fluctuations resulting from bad weather shocks. In all specifications we find strong support for the general Stolper-Samuelson prediction.

2. Why has it been so difficult to link Stolper-Samuelson to the data?

The theoretical and empirical literature on Stolper–Samuelson is immense. Written on the occasion of the 50th anniversary of the theorem, Deardorff (1994) provides a succinct summary of the outcome of decades of theoretical research aimed at extending the proposition beyond the original two-good, two-factor setting.[3] This research agenda has taught us that a move to higher dimensions requires us to separate the content of the theorem in two complementary formulations. The so-called 'Friends and Enemies' version, developed by Ethier (1974) and Jones and Scheinkman (1977), predicts the existence of real factor winners and losers, preserving the 2x2 message of distributional conflict arising from trade. But since this formulation does not identify who the winners and losers are, this approach does not yield a testable hypothesis which one could take to the data. The second formulation, developed by Ethier (1982,1984), postulates a prediction on the effects of the interaction of factor intensities and goods price changes on the direction of nominal factor price changes. The purpose of our paper is to test this falsifiable prediction of the Stolper-Samuelson theorem in an experimental setting that is faithful to the underlying theoretical specification.

The two paramount obstacles to taking the Stolper-Samuelson prediction to the data stem from both the multi-faceted aspects of recent globalization experiences and the complexities of modern economies that are often at odds with the underlying Walrasian structure of the neoclassical model. Globalization often affects economies through other channels besides international trade, like foreign direct investment, migration, outsourcing or technological change resulting from the imports of intermediate products.

[3] Deardorff (1994) is the introductory chapter of the conference volume (Deardorff, Stern et al. (1994)) published on the occasion of the 50th anniversary of the theorem.

A critical assumption behind the Stolper-Samuelson prediction is that goods price changes occur only through exposure to international trade.

The identification problem lies at the core of the trade and wages literature, which has attempted to evaluate whether increased international trade between high- and low-wage economies accounts for the observed increase in the wage premium for skilled manufacturing workers in the US and other OECD countries. Labor economists, such as Katz and Murphy (1992) and Katz and Autor (1999), stress the pre-eminent role of skill-biased technological change rather than in trade in explaining the skill premium. By contrast, Feenstra and Hanson (1999) rationalize the emergence of the skill premium by the increase in outsourcing and trade in intermediate goods.

Since globalisation comes in different forms, the effects of factor prices changes resulting from trade-induced goods price changes might be small. Since historically imports from low-wage economies have been relatively small, the impacts of expanded trade are expected to be of a second order. [4] However, the recent entrance of China into the world economy has changed the landscape and a recent study by Autor, Dorn and Hanson (2012) has found sizable effects of Chinese imports on local US labor markets.

Arguably the biggest challenge in identifying Stolper-Samuelson effects empirically is the required level of disaggregation. The theoretical framework is about the interaction between the market prices in well-defined goods and factor markets over time. However, long-term price data in modern economies are often incomplete and only available in broad product categories which aggregate products with different factor proportions (see Schott, 2004). Furthermore, Stolper-Samuelson is a general equilibrium prediction about how trade-induced changes in goods prices affect the prices of *all* factors of production: different types of labor, capital and land.[5] Because of the well-known challenges of

---

[4]See Krugman (2008) and Krugman (2000) for a discussion of trade's impact on US labor markets under different degrees of exposure to low-wage foreign competition.

[5] A second thread of the empirical literature introduced by Magee (1980) and pursued further in Beaulieu and Magee (2004) examines whether patterns of lobbying or other efforts to influence protection are consistent with the focus of Stolper-Samuelson on the differential impacts of trade on factor rewards. Within a two-factor framework, this research has not found evidence that factor affiliation accounts for

measuring factor returns for capital and land in the complex economies of the modern era of globalisation, the empirical trade literature has focused primarily on explaining the effect of trade on the skill premium.

The rich experience of the first globalisation, which accelerated significantly during the last half of the nineteenth century, offers a promising testing ground for the most general formulation of the Stolper-Samuelson theorem. The period experienced a series of dramatic terms-of-trade shocks as trade costs and political barriers to trade fell, even as technological change in many of the countries lagged behind. The circumstances of the 19th century opening up the Japanese economy coupled with the availability of detailed disaggregated data on market prices and input-output relationships allows for an empirical implementation that is faithful to the theoretical specification.

3. Theoretical framework

Consider a Walrasian economy producing *n* goods from a set of *l* factors of production. Our analysis involves two equilibria and it is assumed that all *n* goods are produced in both equilibria. In the initial autarky equilibrium, goods prices are given by the *n*-vector $\mathbf{p^a}=(p_1^a,p_2^a,\ldots p_n^a)$ and equilibrium factor prices by the *l*-vector $\mathbf{w^a}=(w_1^a,w_2^a,\ldots,w_l^a)$. Production techniques are captured by the economy's technology matrix, denoted by $\mathbf{A}=\langle a_{ij}\rangle$, where a_{ij} are the units of factor *i* necessary to produce one unit of good *j*. Under variable input coefficients, the technology matrix will depend on the factor price vector $\mathbf{w^a}$, i.e. $\mathbf{A}=\mathbf{A}(\mathbf{w}^a)$.

Competitive behavior by a mass of producers all operating under constant returns to scale technologies implies that the general equilibrium relationships between goods and factor prices are given by *n* zero profit conditions. The autarky equilibrium is then characterized by:

$$\mathbf{p^a}=\mathbf{w^a A(w^a)} \qquad (1)$$

---

lobbying efforts. However, as pointed out by Deardorff (1994, p. 32), Magee's results only provide indirect evidence against the strict two good, two-factor formulation of the theorem. They may or may not provide evidence against higher dimensional formulations.

Note that this specification is quite general. It imposes no restrictions on dimensionality and does not require that all factors are used in the production of all goods. Some factors can also be specific to any good or subset of goods.

Suppose that the economy experiences a change, or shock, to goods prices. Such a change could occur from the government imposing taxes on domestic goods, or through adjustment of domestic prices to international prices resulting from a significant reduction in the costs of conducting international trade.[6] Since our empirical implementation exploits an opening up episode, the new equilibrium is a trade equilibrium and the economy's price vector under trade is denoted by $\mathbf{p^t}=(p_1^t,p_2^t,\ldots,p_n^t)$.[7]

A critical assumption of the theory is that the exogenous shock to goods prices does not affect domestic technological conditions. However, since the change in goods prices will change factor prices, the new equilibrium factor price vector will change to $\mathbf{w^t}=(w_1^t,w_2^t,\ldots,w_l^t)$. The trading equilibrium is then characterized by a second set of zero profit conditions:

$$\mathbf{p^t}= \mathbf{w^t A(w^t)} \qquad (2)$$

The Stolper-Samuelson Theorem yields a prediction on how factor prices in the *l* factor markets must respond to changes in the *n* goods prices. Following Ethier (1982), we define a function $b(\mathbf{w})=\mathbf{wA}(\mathbf{w})(\mathbf{p^t}-\mathbf{p^a})$. Applying the mean value theorem, there exists some vector $\hat{\mathbf{w}}$ of factor rewards such that[8]:

$$b(\mathbf{w^t}) = b(\mathbf{w^a}) +(\mathbf{w^t}-\mathbf{w^a})\, db(\hat{\mathbf{w}}), \qquad (3)$$

where $db(\hat{\mathbf{w}})= [\mathbf{A}(\hat{\mathbf{w}})+\hat{\mathbf{w}}\mathbf{dA}(\hat{\mathbf{w}})]\,(\mathbf{p^t}-\mathbf{p^a})$ is the Hessian matrix of b. Since cost minimization implies that $\hat{\mathbf{w}}\mathbf{dA}(\hat{\mathbf{w}})=0$, we obtain

---

[6]In the original formulation of Stolper and Samuelson (1941), changes in goods prices come about because of the imposition of import tariffs.

[7] Because of transportation costs, the domestic price vector $\mathbf{p^t}$ might differ from world prices.

[8] The expression (3) is also the Taylor series expansion of $b(\mathbf{w})$ around $\mathbf{w^a}$.

$$b(\mathbf{w}^t)-b(\mathbf{w}^a) =(\mathbf{w}^t-\mathbf{w}^a)\mathbf{A}(\hat{\mathbf{w}})(\mathbf{p}^t-\mathbf{p}^a) \quad (4)$$

Recognizing that $b(\mathbf{w}^t)-b(\mathbf{w}^a) =(\mathbf{p}^t-\mathbf{p}^a)(\mathbf{p}^t-\mathbf{p}^a)>0$, we obtain a three-way relationship between changes in goods prices, $\Delta\mathbf{p}= \mathbf{p}^t-\mathbf{p}^a$, changes in factor prices, $\Delta\mathbf{w}= \mathbf{w}^t-\mathbf{w}^a$, and the corresponding matrix $\mathbf{A}(\hat{\mathbf{w}})$ of factor input requirements.

Stolper-Samuelson Theorem

Assume an economy experiences a change in goods prices, ceteris paribus. Then the vector of goods price changes, $\Delta\boldsymbol{p}= \boldsymbol{p}^t-\boldsymbol{p}^a$, and the technology matrix $\boldsymbol{A}(\hat{\boldsymbol{w}})$ impose the following restriction on the corresponding change in the factor price vector $\Delta\boldsymbol{w} = \boldsymbol{w}^t-\boldsymbol{w}^a$:

$$\Delta\mathbf{w}\ \mathbf{A}(\hat{\mathbf{w}})\ \Delta\mathbf{p} = \sum_{i=1}^{l} \Delta w_i \left(\sum_{j=1}^{n} a_{ij}\Delta p_j\right) > 0. \quad (5)$$

The Stolper-Samuelson theorem (5) is a prediction on how the vector of *all* factor prices responds to changes in goods prices. The vector on the interaction between factor intensities and changes in all goods prices, $\mathbf{A}(\hat{\mathbf{w}})\ \Delta\mathbf{p}$, imposes a refutable restriction on the vector of factor price changes, $\Delta\mathbf{w}$. It can be interpreted as saying that on average, high values of Δw_i are associated with high values of both a_{ij} and Δp_j. For goods experiencing a sharp increase in their price, factors employed most intensively in the production of them will have a tendency to rise the most in response. The prices of factors used intensively in the production of goods experiencing a decline in their price will have a tendency to fall. Small price changes or less intensive use of a factor will have less of an impact on factor price changes.[9]

[9]This tendency relationship is often referred to as the correlation version of Stolper-Samuelson (see Deardorff (1994)). A shortcoming of the correlation terminology is that it masks the fact that the theorem makes a causal prediction of how factor prices will respond to exogenous changes in goods prices.

An attractive feature of (5) is its great generality. It does not require any special assumptions on dimensionality, technology or complete factor mobility.[10] However, as a trade-off, it does not provide any inference for how changes in good prices affect real factor returns. Under the stated assumption of variable input coefficients the factor input requirements are evaluated at some factor price vector $\hat{\mathbf{w}}$ which could be different from the observed factor prices in either equilibrium. However, the assumption of variable input coefficients is not central to the logic of the Stolper-Samuelson theorem. In fact, Stolper and Samuelson's (1941) original formulation was formulated under fixed input coefficients. Since changes in relative factor prices most often result in organizational changes that modify entire production relationships rather than leading to a 'movements' along technological isoquants, it appears justified to assume that the economy's input coefficients are fixed. So for the rest of the paper we assume that **A** is independent of factor prices and therefore the same in both equilibria.[11]

The prediction (5) is a clear generalization of the familiar two-good, two-factor formulation and can be illustrated in Figure 1. Under fixed input coefficients, the zero profit conditions (1) define two lines I and II with the intersection E^a determining the factor price vector (w_1^a, w_2^a) in the initial equilibrium. For illustrative purposes, assume that the factors are different types of labor: factor 1 is low skilled and factor 2 is high skilled.

[Insert Figure 1 here]

Figure 1 is drawn such that the production of good 2 is relatively intensive in the low-skilled factor 1, i.e. $(a_{12}/a_{22})>(a_{11}/a_{21})$. Assume now that the economy experiences an exogenous increase in the price of the high skill-intensive good 1, denoted by $\Delta p_1>0$, while the price of the low-skill intensive good 2 remains the same. This will lead to an

[10] Since the derivation depends upon cost minimization (and the applicability of the mean value theorem), the result is also invariant with respect to changes in endowments. There are also no restrictions on the production path between the two equilibria (see Ethier (1982, p. 339)).

[11] The assumption of fixed input coefficients also implies that b(w) is continuously differentiable over some path of non-negative factor prices, which justifies the application of the mean value theorem to b(w). Furthermore, our technology matrix also fulfils the assumption of no joint production

upward shift of good 1's zero profit line from I to I', while good 2's zero profit line remains at II. The direction of the resulting change in factor prices is determined by the relative magnitude of the factor input requirements. To accommodate the increase in the production of the high-skill intensive good 1 requires a relatively larger amount of the high skill-intensive factor 2 than good 2 is able to release of it. However, since the price of good 2 has remained unchanged, this can only be accomplished by an increase in the price of factor 2 and a decrease in the price of factor 1.

The Stolper-Samuelson logic has been applied to investigate the role of international trade on factor returns; in particular to account for the observed increase in the skill-premium. The difficulty of applying Stolper-Samuelson to account for this observation is that the Stolper-Samuelson theorem assumes trade-induced goods price changes in the absence of technological change. The identification problem that arises in the presence of technological change can be illustrated in Figure 1. Specifically, an increase in the observed skill-premium, i.e. $w_2^t/w_1^t > w_2^a/w_1^a$, could have been also the result of skill-based technological change, i.e. a decline in a_{21} , assuming for simplicity that a_{11}/a_{21} has remained unchanged. In the presence technological change, the effects of trade-induced changes in goods prices on factor prices are difficult to identify.

4. Empirical strategy: Using a natural experiment to test Stolper-Samuelson.

Our research design exploits a natural experiment to establish the empirical validity of the Stolper-Samuelson prediction (5) derived in section 3. As Heckman (2000, pp. 84-85) notes, natural experiments have been used widely in empirical economics as an instrumenting strategy for finding estimates of what he labels "causal parameters." Our strategy employs a natural experiment in a different way; it tests a general theoretical prediction about the response of factor prices in a competitive economy to an externally-imposed shock to the vector of prices prevailing in the economy. The natural experiment in this case was brought about by the abrupt abandonment of autarky that followed in the wake of Admiral Perry's 1853 expedition to Japan and the 1858 Treaty of Amity and Commerce, but the results are not specific to that episode. Provided the other criteria of the research design are met— the *ceteris paribus* assumptions hold and the data are reasonably closely matched with the

theoretical constructs—the results apply in principle to the economies of other countries and other time periods as long as they conform reasonably well to the assumptions of a Walrasian economy.[12] McCloskey (1976, p. 447) argues that the economies of the late 18th and 19th centuries— with their atomistic competition— most likely come closer to the assumptions of a Walrasian economy than the economies of the 20th or 21st centuries.

The nature of the opening up and our choice of experimental windows ensure a testing environment that fulfills the critical *ceteris paribus* assumptions of the theory. A key advantage of the data set employed in this study is that it conforms well to the neo-classical assumptions of the model. Price data are for the most part transactions prices for the same commodity over the test periods. The technology matrix is based entirely upon quantitative data on factor requirements at the disaggregated product level and measured in the same units (days of labor, acres of land) for which factor prices are available.

Japan as a neoclassical market economy

An attractive feature of this natural experiment is that Japan's economy during the middle of the nineteenth century can be characterized as a Walrasian market economy. Markets for goods were competitive, products were homogenous, technologies were for the most part constant returns to scale and factor prices were the outcome of market forces. Bernhofen and Brown (2004, p. 56) and Bernhofen and Brown (2005, p. 214) provide evidence that Japanese goods markets during the period of the natural experiment were highly competitive. During the autarky period, the Japanese economy produced the full range of agricultural and industrial goods consumed by its population, albeit with a significant degree of regional specialization.[13] By the late autarky period, final and intermediate products were being traded in markets throughout the country; the most

[12] Or in the terms proposed by Roe and Just (2009), our formal theoretical model and the characterization of the Japanese economy as being representative of *any* Walrasian economy will confer our test results external validity. The discussion below draws upon two other ideas implicit in the empirical trade literature that Roe and Just (2009) articulate: data suitability (construct validity) and the potential for causal inference within the terms of the theoretical model (internal validity).

[13] The main exceptions are small imports of woolens from the west and ginseng, woven silk and sugar from the east. These imports had dwindled to insignificance by the 1840s.

important market remained the commercial center of the country, Osaka. In addition, the immense administrative center of Edo (present-day Tokyo) offered one of the largest agglomerations of consumers on the planet. The high degree of commercialization accounts for the rich deposits of price records stretching back into the eighteenth century for most of the commodities that eventually entered into trade.

The bulk of production, including the products that entered into trade once the economy opened up in 1859, met the strict definition of homogenous commodities (those that were traded on exchanges) found in Rauch (1999). Price quotations for intermediate products such as raw cotton, cotton yarn, raw silk, Japanese wax (used for candles) and bar iron as well as final goods such as bleached cotton cloth, rice and soybeans used in this study are available from the Osaka market (the largest in the country) for autarky and the first twenty years of open trade.

Production technologies in use during the period reflected constant returns to scale. With the exception of the sake and soya sauce brewing industries and a few of the larger copper mines, production took place in the countryside in the farmer households that made up 80 percent of the population or in small firms of 5 to 20 workers. The batch processes then in use limited the scale economies. The most important branches of the economy producing finished manufactured goods (silk, cotton, iron manufactures) were vertically dis-integrated and linked by a dense network of markets for intermediate goods and raw materials.

Markets for labor and land responded to price signals to reallocate productive resources. Despite the presence of formal restrictions on occupational changes, the growth of by-employment opportunities outside of the peasant household generated an increase in an active labor market. Saitō (2009, p. 186) notes that by the first half of the nineteenth century there was "a well-integrated labor market between the peasant farm household and non-farm sectors within a regional setting." Nishikawa (1978) uses data from a detailed survey of the Chōshū domain in the 1840s to demonstrate that wages were close to their marginal product. In addition, the concentration of production of both agricultural and industrial goods in multi-product peasant household firms meant that reallocation of labor to alternative activities could take place without significant

migration.[14] This stands in stark contrast to labor market rigidities often found in modern developing economies which have resulted in limited potential for sectoral labor reallocation, as emphasized in Goldberg and Pavcnik (2007).

Saitō (2009) and Saito (2010, pp. 247-252) document the ways in which markets for land functioned during the middle third of the nineteenth century. Although complete ownership rights for land could not be legally transferred, a vigorous market in short-term leases and a pawning market in land allowed the reallocation of land to alternative use. Records of villages and landlords provide ample evidence of both rents and prices for dryfields and paddyland for the period of our study.

Opening up and price shocks

From near autarky at the beginning of 1859, Japan moved to a very open trade regime *de jure* by the end of that year, which became *de facto* in 1864 after western military interventions convinced the Tokugawa rulers of Japan to fully meet the obligations to open markets stated in the trade treaties of commerce with western powers. The opening up exerted a price shock to the Japanese economy that was so severe that it is known in the Japanese historical literature as a price revolution (see Miyamoto (1980)).

[Insert Figure 2 here]

Figure 2 shows the main events that marked the twenty-nine years spanning the last years of autarky and the first decades of open trade. Admiral Perry arrived in Tokyo harbor in July of 1853 and demanded that Japan be willing to provide safe harbor for whalers whose ships were disabled. By 1858, trade treaties had been negotiated between Japan and several western powers that led to the opening up of trade on July 4, 1859. Japan's raw silk markets were not fully opened up to trade until late 1864, when the Shogunate abandoned attempts to control the amount of silk coming on to the market in the face of the western shelling of shore batteries at Shimonoseki. The Meiji revolution of

[14] See Tanimoto (1998) for a discussion of this kind of reallocation in Nagano Prefecture, from activities related to cotton cloth production to sericulture.

March, 1868 and the most severe harvest failures since the Tenpō famine of the 1830s marked the late 1860s.

[Insert Figure 3 here]

Figure 3 shows the evolution of the monthly price of the main export (raw silk) relative to the largest import, cotton yarn, on the Osaka market during the 30 year time period of 1850-1879. Because of the monetary changes brought about by fiscal challenges and the integration of the Japanese gold-based system with the international monetary system, nominal prices almost tripled in the decade from 1860 to 1870. This initial monetary and trade shock helps to account for the increase in volatility in the relative price compared to the previous decade. The disruption to cotton markets of the American Civil War in the first half of the 1860s and the harvest failures of 1869 were additional factors that accounted for the severe swings in in relative prices during the 1860s. After 1870 relative prices stabilized and the period from 1871-79 appears to reflect price conditions in a 'long-term trade equilibrium'. Comparing the ratio of the price of silk to yarn in the 'autarky equilibrium (1850-1857)' to the ratio under free trade, we find an increase in relative price from 1.63 to 2.57 in the early 1870s (1871-75) and to 2.59 in the late 1870s (1876-1879).[15]

Given Japan's commercial and physical geography, the treaty ports designated by the trade treaties ensured that the price shocks penetrated deep into the interior. Japan's economy was primarily oriented towards the Pacific Ocean. Even during autarky, the most productive part of the economy and much of the population was found in three coastal plains bordered by steep mountains that were adjacent to the Pacific Ocean or the Inland Sea: the Kansai region (centered on Kyoto and Osaka), the Kantō region (centered on Edo) and the Tōkai region (located midway between Osaka and Tokyo and centered on Nagoya). The two main treaty ports of Hiogo (later Kobe) and Kanagawa (later Yokohama) were established in close proximity to the Kansai and the Kantō regions; both provided ready access by coastal transportation to the Tōkai region. In addition, the

[15] Note that this is an increase in the barter terms of trade between silk and yarn of 58 to 69 percent over the period.

cost of transporting the most popular imports (cotton textiles and woolen cloth) to the interior was a relatively small share of its delivered price. Likewise, the chief exports of silk and tea had a high value to weight ratio and could be transported from the interior at a small share of the FOB price.[16] Low import tariffs equivalent to less than five percent *ad valorem* imposed by the western powers also ensured that import protection would not cushion the impact of the price shock on Japan's markets.

Experimental windows and the price shock

The length of the time interval separating the 'autarky equilibrium' from the 'trade equilibrium' for a test of the prediction (5) involves a critical trade-off. The test interval must be long enough to allow for factor price changes to signal a reallocation of factors across sectors and to allow for the reallocation of the resources in response to the initial price shock (or shocks) to take place. In addition, a longer test interval ensures that decision-makers will be able to distinguish longer-term permanent price shocks from short-term temporary shocks. At the same time, the longer the interval, the more likely that the underlying *ceteris paribus* conditions are not met: tastes or technologies may change in significant ways.

The circumstances and historical timing of Japan's opening up allow us to observe experimental windows of autarky and open trade that meet the two criteria: a long enough time period between autarky and trade for reallocation to take place, but at the same time a period short enough so that the ceteris paribus conditions are fulfilled. The sequence of historical events (Figure 2) and price swings (Figure 3) suggest that the early to mid-1850s (1850-57) are best suited for the autarky window. For the open trade window, we consider two candidates: 1871-1875 and 1876-1879.

Consider first the time for reallocation. The agricultural technologies used in Japan's two main export industries, tea and sericulture, imply significant lags in the ability of producers to respond to an initial price shock. Syrski (1872, p. 211) notes that

[16]For example, Syrski (1872, p. 261) notes that the cost of transportation from the producing region to the export port of Yokohama was only 3 percent of the farmgate price of silkworm eggs.

tea shrubs required at least four years of growth before they were ready to be harvested. Syrski (1872) reports that mulberry trees, which provide the leaves that are fed to silkworms, required from three to up to eight years to fully mature. Some of the skills used in the sericulture industry also required several years to develop. Only one day may be needed to train a family member to prepare a sheet of silkworm eggs; other tasks such as picking mulberry leaves may require two years to master. A well-known handbook of the sericulture industry estimated that learning how to reel silk (draw silk from a cocoon) could take up to the 25 years to master.[17]

Although goods markets were as well-integrated as possible given the availability of relatively inexpensive water transport, labor markets and land markets (particularly rents on land) may have responded only slowly in response to factor prices. As will become apparent with the presentation of the results, adjustment to the relative price shock evident in Figure 3 most likely required a minimum of 10 years and more likely close to 15 years.

Experimental windows and the ceteris paribus conditions

Although the transition period from autarky to open trade exposed the Japanese to new encounters with western goods and steam-based technologies, both preferences and production technologies remained stable during the early trade period. Despite the enthusiasm of some elites for western culture and institutions after the Meiji restoration of 1868, preferences for consumption remained largely unchanged during the first two decades of open trade. The imports most sensitive to changes in tastes would have been imports of cotton and woolen textiles, which accounted for one-half of Japan's imports through 1880. Uchida (1988) and Tamura (2001) provide ample evidence that western merchants quickly learned about the need to adapt western cloths to Japanese styles of clothing and tastes. The Japanese readily incorporated imported machine-spun yarns in

[17] See Narita, Koyana et al. (1978, pp. 7-8). This source is a treatise on sericulture that was first published in 1814.

their methods of weaving with traditional narrow handlooms, but the cloths produced adhered to the 14-inch widths required for traditional Japanese clothing.[18]

During the first two decades of open trade, virtually all of the production of tradable goods continued to rely upon the traditional technologies in use during the autarky period. The adoption of steam-based technologies from the west was limited to a few sites. High interest rates relative to the price of labor was one important reason for the persistence of labor-intensive methods. Anonymous (1875, p. 90) notes nominal rates of 18 percent ca. 1874 made mechanization of silk reeling using the steam technologies found in France or Italy unprofitable. Lyman (1879, pp. 258-259) argues that drilling oil wells with hand labor was lower-cost than using steam technology and Gribble (1874, p. 99) makes a similar argument for the continued use of physical labor to produce vegetable (or Japan) wax rather than western-manufactured steam presses.

By the close of the test period (1879), the Meiji government and some western investors had introduced some western methods in government-run armories and a handful of coal and copper mines, but most steps in the processes relied upon traditional Japanese technologies.[19] Land transportation continued to rely upon pack horses. The largest change occurred in coastal shipping, with the introduction of some steamships and western sailing vessels.[20]

---

[18] See German Consul in Hiogo (1873). The German Consul in Edo (1873) reports that the use of woolen cloth for western dress was restricted to government uniforms of various kinds and some of the wealthiest classes.

[19] British consuls in the treaty ports complained repeatedly that the government's hostility to foreign ownership prevented modernization of the mining and metallurgy sectors. See the summary report by Plunkett (1875). Full-scale adoption of western methods in the mining sector took place during the 1880s.

[20] Unlike in Latin America, the Anglo-Saxon regions of recent settlement or colonial Africa, foreign capital played virtually no role in the establishment and construction of modern transportation systems in Japan.

5. Data

The test of the Stolper-Samuelson prediction (5) requires matching data on goods prices, factor input requirements and factor prices. The full details of the dataset utilized in this study are available in a Data Appendix. This discussion highlights the key issues associated with collecting data on each of the three components and the approach that was followed to address the substantial price inflation that took place over the test periods.

Product scope and prices

Our empirical study uses price and factor input requirements for 18 traded goods. Nine of these were exports and nine were imports. Although in principle the test of the theorem applies to any subset of goods markets, a test that incorporates as wide a range of goods as possible is most preferable. Table 1 lists the goods for which price data were available, the change in price over the two test intervals and two measures of the importance of trade in the good.

[Insert Table 1 here]

All prices are expressed in terms of gold en (one-thousandth of a gold ryō or yen) per kin or catty, which was equivalent to 1.33 pounds. [21] To ensure that the inflation that occurred from the 1850s through the early 1870s does not distort the results, the price index for non-trade goods compiled by Shinbo (1978, Table 5-10) was used to convert prices to constant *en* using 1855 as the base year.

As may be expected, the value of raw silk and silkworm eggs was about ten times the value of cotton yarn. For the most part, prices of imports tended to fall and prices of exports tended to rise after the opening up.[22] The steep increase in the price of silkworm

[21] The ryō was the gold-based currency of Tokugawa Japan. It was replaced with the yen at a ratio of one to one in 1871.Prices in Osaka were quoted in silver *monme*. They were converted to ryō using the monthly exchange rates found in Miyamoto and Ōsaka Daigaku Kinsei Bukkashi Kenkyūkai (1963). After 1871, the *rin* was the equivalent of the *en* and equal to one-thousandth of a yen.

[22] In our higher dimensional formulation, there is no direct prediction on how the direction of price changes from autarky to open trade affect the direction of trade for each individual commodity.

eggs resulted from strong demand from producers in France and Italy, where a large share of domestic silkworms were infected with the pébrine disease. Because of its isolation, Japan's silkworm population escaped the impact of this disease, which devastated European sericulture through the mid-1870s. The decline in the prices of cotton goods reflected the impact of both inexpensive sources of supply for raw cotton outside of Japan and the immense productivity of the British cotton textile industry.

A quick look at the data on the shares of trade and of domestic consumption (or production) in columns 5 and 6 confirms the importance of cotton goods for the tradable sector of the Japanese economy. The other sectors most strongly influenced by imports were sugar and iron manufacture. Japan imported sugar from China and Formosa and imported iron manufactures from Great Britain and Belgium.[23] Overall, the eighteen commodities analyzed in this test of Stolper-Samuelson accounted for about 44 percent of imports and three-quarters of exports.

As was typical for virtually all developing countries and colonies during the 19th and early 20th centuries, Japan's export sector was highly specialized. All products of the sericulture sector, of which raw silk and silkworm eggs were most important, accounted for about one-half of Japanese exports. About one-half of the sector's products were exported. Green tea accounted for another one-quarter of exports and a very large share of the annual production was exported. Copper would later grow in importance; most of the exports were from a few mines and smelters that had begun to employ western steam technology for conveyance of iron ore and for powering pumps.

An important feature of the price data used here is that virtually all of it is from continuous series of monthly or annual price quotes for standard qualities of commodities traded on the Osaka market, which was the largest market in Japan. Compiled by Miyamoto and Ōsaka Daigaku Kinsei Bukkashi Kenkyūkai (1963), the data are available up through the end of 1879 (the close of our test periods).The most important exceptions

[23] Reports from western observers and Japanese iron manufacturers suggest that production of bar iron, which used smelted iron sand from the traditional *tatara* furnace, almost ceased during the 1870s in the face of western imports.

are the prices of silkworm eggs, which are from merchant records and consular reports, and the prices of camphor, seaweed and cuttlefish. We exploit the availability of the high frequency price data in the robustness checks for the key Stolper-Samuelson results.

Table 1 includes the prices of yarn and cotton cloth, which amounted to about one-third of domestic Japanese consumption (by weight) by the late 1870s and constituted the most important import by value. It also includes the price of bar iron, which was determinant for the prices of manufactured iron products such as nails. Japan's imports of iron accounted for about one-quarter of its domestic consumption by the end of the 1870s. Brown and white sugar were the most important import of foodstuffs, once the effects of the poor harvests of 1868 and 1869 had been overcome. Silk and silkworm eggs at times accounted for three-quarters of the value of Japanese exports. Tea was the third most important export, eventually followed by copper.

Matrix of input requirements

The second data component is the **A** matrix for the eighteen commodities for which price information spanning the periods of the test could be found. A data appendix available from the authors provides more detail on the input requirements found in the A matrix and a discussion of the main sources used for it. The core source for agriculturally-based commodities (rice, soybeans, sugarcane, tea, cotton, silkworm eggs and raw silk) was the multi-volume survey of agricultural production conditions from the mid-1880s known as the NōjiChōsa (see Chō, Shōda et al. (1979)). This source details the labor, capital and land requirements for all of the tasks associated with producing virtually all agricultural commodities at the prefectural level. For goods produced throughout Japan (soy and rice), the averages were used for resource requirements. For goods where regional specialization mattered (sugarcane, cotton, tea, and products of sericulture), only those prefectures contributing significantly to production at the time of the Meiji census of production in 1874 were included.[24] To ensure the appropriate

[24] See Le Gendre (1878) for an English-language summary of the main results.

estimate of productivity in silk reeling, only estimates from prefectures still practicing hand reeling were used. In addition, the data for tea was restricted to varieties (such as *sencha*) that would likely be exported.

The production conditions of these industries suggested that a division of labor into three categories (skilled male, unskilled male and female) provided an appropriate compromise between detail and comprehensive coverage. Skilled male labor played an important role at certain points in Japanese production processes (in processing tea or scutching cotton, for example). It was also important for several processes in mining and metallurgy. Since a large share of production was by household units based on the farm, female labor was ubiquitous in Japan.[25] Male labor was rarely involved with tasks such as picking tea, reeling silk, spinning, weaving or preparing copper ore for roasting. For tasks that may have involved the labor of male and female labor, the ratios in the respective sources were used. In addition, the **A** matrix took account of three important intermediate inputs: fish fertilizer, which was produced using herring caught off the coast of Hokkaido, charcoal, which was used in enormous quantities in metallurgy, and wood.

Data for the remaining industries drew upon a plethora of Japanese and western sources. Uemura (1986) and Oka and Yamazaki (1983) are key sources for sugar refining. Tanimoto (1998) provides an excellent review of the technologies in use in cotton processing, spinning and weaving. Doi (1984) and Takahashi (1995) provide excellent coverage of the tatara technology used in the production of iron. Lyman (1879) is an oft-cited source for details of iron production technology and copper mining and smelting.

Factor prices

The final component is the vector of factor prices for the three test periods. For evidence on wages that spans the entire period, Saitō (1973) and Saitō (1998) (a summary volume) provide series of wages for male and female workers in eastern (Kantō) and western (Kinai) Japan. With the exception of 1850-1857, the five-year

[25] This study has included the labor of older children with the labor of women.

averages found in this source are used for unskilled male workers and for female works. For the first (autarky) period, the individual wage observations found in Saitō (1973) and a few other sources were used. The data have sufficient regional coverage to estimate wages for both all of Japan and for the two separate regions.

In calculating the capital requirements for the A matrix, the user cost of capital calculated in constant ryō of 1851-1855 was used. For this reason, capital is treated as the numéraire factor for the autarky period and is priced at 1 in 1851-55. The most important influence on the relative price of capital that could have changed over the test period is the interest rate. Anonymous (1875, p. 90) notes nominal rates of 18 percent ca. 1874 in a discussion of the potential for mechanizing the silk reeling industry, which is consistent with the observation of Lyman (1879) that high rates of interest prevailing in Japan in the late 1870s were not conducive to capital investments. Saitō and Settsu (2006) argue that interest rates of 12 to 15 percent were most common during the late autarky period, which would suggest that the price of capital rose after the opening up.[26] Depreciation rates on Japanese capital such as the tatara furnace used in the smelting of iron sand were about ten percent. For wooden tools or wooden boats, which are more likely closer to the kinds of capital employed in the tradable sector, they were closer to 20 percent. At this rate of depreciation and assuming a 15 percent interest rate in autarky, a lower bound estimate of the rise in the price of capital would be 0.085. This estimate is used in the results presented below,

The final factor of production is land. With the exception of rice, every other agriculturally-based commodity analysed in this study was produced on “dry fields” rather than the paddies used for rice production. Fortunately, detailed cross-sectional data are available on rents and values for this kind of land both for the Kansai and the Kantō regions for the late autarky and free trade period. Hedonic regressions of the data in these

[26] Mitchener and Ohnuki (2007, Figure 1) report interest rates for the core Kantō, Kinai and Chugoku regions in the range of 11 to 13 percent. At these high rates of depreciation, a large change in the interest rate is needed to achieve even a modest reduction in the user cost of capital.

sources and other data established base rents for dry fields of a "standard" quality.[27] Additional time series sources are available for both regions to provide guidance on the trend in rents through the early 1870s. Additional detailed data on land values from sales within the same district are available from the *Japanese Economist* for 1875 and 1877. Hedonic regressions of these land values were used for estimates of rents during the second half of the 1870s.

6. Empirical implementation

We test the Stolper-Samuelson prediction (5) in the context of the natural experimental setting described in section 4. Our choice of experimental windows was guided by the objective of observing economy wide factor price responses to trade-induced goods price changes, *ceteris paribus*. The tests of the hypothesis use both experimental windows in a benchmark case, which assumes that land was essentially undifferentiated across Japan. Our examination of robustness of the initial test results allows for differentiated responses of land values in east and west Japan to the price shock of opening up. It is possible that mean prices may not fully capture the relevant features of the distribution of product and factor prices. For that reason, we implement a test for robustness of the results that allows us to assign a level of significance to the hypothesis test results.

Benchmark case: Fully integrated factor markets

Our benchmark specification assumes that all factor markets were nationally integrated, including the market for land. The results are reported in Panel A of Table 2, which provides the two subcalculations and the outcome for the calculation of the inner

[27] The data from Kinai region includes a dataset on 33 plots across three villages, a dataset on rental prices for 20-25 plots that were rented out over about 40 years, the rents charged in a village for two different qualities of land and average sales prices for a mix of paddy and dry land fields. All of the data are from the area around Osaka. The data from the Kantō are from a silk-producing region north of Tokyo (see Waseda Daigaku Keizaishi Gakkai (1960)) (and include observations through the early 1850s) and a data set of land transactions from the Gunma prefecture, which was also a silk producing area further north and east of Tokyo (see Gunma-ken (1977)). The data on sales prices were converted to rents by multiplying the sales prices by the 9.6 percent used in the Japanese valuation of land at the time.

product for the two free trade experimental windows of (1871-1875) and (1876-1879). The second and fourth columns provide the results of the calculation of $\sum_j a_{ij}\Delta p_j$; each row has the results for one of the five factors. The third and fifth columns provide the changes in factor prices (Δw_i) over the test intervals. The final row in the table has the resulting inner product of the restriction captured by $\mathbf{A}\Delta\mathbf{p}$ and the actual change in factor prices $\Delta\mathbf{w}$ for the two test intervals.

[Insert Table 2 here]

The direction of the changes in the factor prices differ over the two test periods. The decline in wages ranged from four percent for skilled male labor to fourteen percent for female labor. An increase in interest rates drove the rise in the price capital. Finally, the weighted average of land rents fell 13 percent. Wages showed some recovery one percent for female workers up to 30 percent (relative to 1850-1857) for unskilled male labor. The rent on land shows a (very) modest recovery to only a decline of three percent relative to the 1850s. As we will see in the section below, the average for Japan obscures substantial inter-regional differences in the behavior of rents.

Looking at the components of the inner product of product prices and factor intensities reveals that unskilled male labor is much more important than female or skilled labor in influencing the inner product. With a change in unit price that is substantially larger than the change in the wage paid unskilled labor, capital has a strong potential to play an important role as well. Land could be viewed as the third most important factor in influencing the size of the test statistic. The final role of Panel A confirms that for both periods, the inner product $\Delta\mathbf{w}(\mathbf{A}\Delta\mathbf{p})$ is positive and the theoretical prediction of the Stolper-Samuelson theorem is confirmed.

As an alternative to the approach taken in Table 2, we can express the inner product by calculating the changes in factor prices weighted by the factor requirements ($\sum_i \boldsymbol{a_{ij}}\Delta w_i$.) and compare this to the price change Δp_i for each of the 18 goods. Figure 4 illustrates the results of these calculations for each of the two test intervals. The goods price change is on the horizontal axis (measured in one-tenth of a yen or ryō to facilitate comparison) and the corresponding changes in factor intensity-weighted factor prices on

the vertical axis. Imports are indicated by a triangle and exports are represented by a circle. The relative importance of either traded product in Japan's trade is indicated by the size of the marker. As the graphs for both periods indicate, the changes in factor prices for goods in sericulture were for the most part strongly correlated with the relative factor intensities. The resulting values of $\sum_i \boldsymbol{a_{ij}} \Delta w_i$ occurred at the same time that there was a steep rise in price of sericulture products *per kin*. None of the other products played such an important role in trade, nor did they require anywhere near the amount of factor services as products in sericulture.

[Insert Figure 4 here]

It is important not to be tempted to interpret raw silk and silkworm eggs as outliers, or even potential coding errors, in determining the positive relationship suggested by the theory. The 18 data points in Figure 4 do not represent 18 repeated observations of a bivariate relationship, but represent the outcomes of the effects of an exogenous price shock on 18 interrelated markets. The large magnitudes on "raw silk" and "silkworm eggs" reflects the large comparative advantage the Japanese' economy had in these sectors and constitute therefore an integral component in the overall pattern of price predictions.

Allowing for differences in the rent on land

Although the results appear to be reasonably robust with respect to the test period chosen, it is possible that averages across regions are concealing movements in factor prices that are inconsistent with the test proposition in equation (5). After all, there is a long historiographical tradition that argues that even as the eastern part of Japan (particularly the Kantō region and adjacent silk-producing areas) benefited from the extraordinary rise in the price of silk above levels prevailing in autarky, the western part of Japan centered on the Kansai was hit with the incursion of inexpensive Chinese cotton, cheap machine-made British yarn and cloth and low-cost Formosan sugar. Scattered qualitative evidence found in the records of a landowning family that lived in the Nagano prefecture, which produced the largest amount of silk in Japan, suggests that the expansion of sericulture after the opening up raised rents on most dryfields owned by the

family.[28] Detailed regression analysis of land (lease) prices from the silk-growing Gunma prefecture and Ashikaga county in the silk-raising region near Tokyo provide quantitative evidence that rents on dryfields may have actually risen from the 1850s through the 1870s. The estimated increase in real land prices in western Japan contrasts with the decline in rents found in the Kansai. Since very little silk for export was produced in western Japan and most other products were produced in the western part of the country, it may be more reasonable to differentiate between land in the east (used in sericulture and indigo dye production) and land in the west (used in cotton, sugar and most other products listed in Table 1). Panel B of Table 2 extends the test results presented in Panel A to allow for regionally differential impacts of price changes on land rents. By 1879, dryfield rents had risen about forty percent in the east, and they had fallen about 40 percent in the west.[29]

[Insert Figure 5 here]

The results, which appear in the third and fifth columns of Panel B, are even stronger than was the case under the assumption of equal land rents across all of Japan. The test expression (5) is positive for both test periods. Figure 5 plots the values of the factor prices weighted by factor intensity. The main result from Figure 4 remains. The high factor intensity and large price changes for products of the sericulture industry were consistent with the observed changes in factor prices for both periods. In addition, allowing for regional differences in the evolution of rents means that products of the cotton industry (from raw cotton through cotton cloth) saw changes in factor prices that were consistent with the substantial declines in product prices.

---

[28] See Nakamura (1962).

[29] In this event, the Kansai definition includes the remainder of the west of Japan. Further refinement of the regional definitions is possible, but unlikely to materially affect the results presented here.

Allowing for sampling variability

We would like to assess the extent to which the positive test outcomes in Table 2 are sensitive to sampling error. First, express the test of the Stolper-Samuelson theorem as a hypothesis test subject to Type I error. The null hypothesis is

$$H_o: \sum_{i=1}^{l} \Delta w_{iT} \sum_{j=1}^{n} a_{ij} \Delta p_{jT} \leq 0 \,, \qquad (6)$$

where T is one of the two test intervals. The alternative hypothesis is of course that the inner product is greater zero:

$$H_A: \sum_{i=1}^{l} \Delta w_{iT} \sum_{j=1}^{n} a_{ij} \Delta p_{jT} > 0 \qquad (7)$$

Note that Δw_{iT} is measured with $\sum_{n=1}^{n_1} \widetilde{w}_{i1} - \sum_{n=1}^{n_2} \widetilde{w}_{i2}$. n_1 is the number of observations available for the first part of the test interval (1850-1857) and n_2 is the number observations available in the terminal period of the test interval period (t=2 for 1871-75 and t=3 for 1876-1879). The observed $\widetilde{w}_{i1}$ equals the true w_{i1} (the true wage during the late autarky years of 1850-1857) plus an error term ε_{it}. (Similar notation can be used to express the relationship of the observed change in prices and the true prices.) The core issue is how to assess the impact of sampling error (ε_{it}) on the components Δw_{iT} and Δp_{jT} without making heroic assumptions about underlying distributions.

The opening up of Japan to international trade created an extraordinary natural experiment; unfortunately, some of the data that historians have gleaned from the years that are of most interest are available at only relatively low frequencies (n_t can be small), since the test periods precede systematic collection of wage and price data by governmental authorities. Although Table 1 suggests that this issue would be of concern for the 11 commodities for which only annual (or periodic) data are available, a quick look at Figures 4 and 5 reveals that sampling error would really matter for the price of silkworm eggs; the other prices had at best only a minor influence on the calculated value of the inner product. Sampling variability would matter more for most factor prices. There are only five observations for skilled wages and seven observations for silkworm egg prices in the autarky period, for example. Application of standard t-tests to account

for this sampling variability carries with it an assumption that ε_{it} is normally distributed, which may be untenable.

Fortunately, statisticians have long known that permutation tests can provide an exact level of Type I error ($\boldsymbol{\alpha}$) for situations where inference is based upon relatively small samples coming from unknown distributions.[30] The primary assumption for carrying out such tests is that ε_{it} is random. For the wide variety of factor and product prices used in this study, an assumption of randomness is appropriate. Historical accidents may lead to the preservation of one merchant's or farmer's records and the destruction of another's; why wage surveys in one village survived to be published but were lost in another can be reasonably viewed as a (random) accident of history. The price records coming from the Osaka market could have just as well been sampled from midweek rather than at the beginning or end of the month. The assumption of randomness is surely less stringent than asserting that a particular distribution underlies the data that are available from historical sources. The results of the tests, which are performed with Monte Carlo simulations, offer an asymptotic Type I error ($\boldsymbol{\alpha}$) for each inner product presented in Table 2. The results are conditional only on the sample; given the sample, they are unconditional with respect to the underlying distribution of the population.

Appendix One provides more details on the implementation of the permutation tests of the null hypothesis. Essentially, the null hypothesis is that the period from which prices are taken (autarky or open trade, in this case) has no significant influence on the value of the inner product. In its simplest form, the test calculates all possible permutations of "pre" and "post" samples of size n_1 and n_2 from the n_1+n_2 values of prices available for a particular product or factor; the change in price is calculated for each of the permutations and the resulting inner product is calculated. The proportion of inner products calculated under the assumption of the null hypothesis that is greater than

[30] See Ludbrook and Dudley (1998) and Ernst (2004) for useful general introduction and Appendix 2 for a discussion of how the tests were implemented. We are grateful to Gbeton Somasse for his assistance in both identifying the permutation test approach and carrying out the simulations required to implement it.

the test values appearing in Table 2 is the probability of Type I error. For those situations where the number of permutations exceeds 10,000, the permutations used in the calculation are drawn randomly from the population of potential permutations in Monte Carlo simulations.

[Insert Table 3 here]

Table 3 presents the results of the permutation tests for the two test intervals and under the assumptions of regionally undifferentiated and differentiated land. Two values of $\boldsymbol{\alpha}$ are presented for each case. The first excludes rents from the permutation exercise under the assumption that the values calculated from the hedonic regressions are measured without error, and the second includes rents in the exercise. Overall, Type I error ranges from a high of 18 percent for the first test period (including rents) down to 0.1 to 0.4 percent during the second test period (excluding rents). The test of the full general equilibrium implications of the Stolper-Samuelson Theorem is robust to sampling error.

7. Conclusions

The Stolper-Samuelson theorem is one of the central propositions of general equilibrium (trade) theory. Half a century of theoretical research efforts aimed at generalizing Stolper and Samuelson's (1941) two-good, two-factor formulation has resulted in a refutable "general equilibrium systems prediction" that can be derived under quite general conditions. This paper has broken new ground by collecting the data around a natural experimental setting that allowed us to conduct the first scientific test of this fundamental proposition. We can with considerable confidence assert that the theorem holds.

We conclude by summarizing the two key messages of our paper. First, the empirical domain of Stolper-Samuelson is beyond the well-known discussion about the rise of the skill-premium. Specifically, it pertains to the response of all factors of production, including capital and land, to globalization. Second, our empirical verification of the systems formulation of Stolper-Samuelson should not be only of

interest to international economists but to all who care about the empirical verification of Walrasian general equilibrium theory.

References

Ahlfeld, G., S. J. Redding, et al. (2012). The economics of density: evidence from the Berlin Wall. *CEP Discussion Paper*. London, London School of Economics.

Anonymous (1875). "Über den gegenwärtigen Stand des Seidenhandels in Japan und die Schwierigkeiten, mit denen derselbe zu kämpfen hat." *Preussisches Handelsarchiv*(33): 189-190.

Atkinson, R. W. (1881). "The Chemistry of Sake Brewing." *Memoirs of the Science Department Tokio Daigaku (University of Tokyo)* (No. 6).

Autor, D. H., D. Dorn, et al. (2012). The China Syndrome: Local Labor Market Effects of Import Competition in the United States, National Bureau of Economic Research, Inc, NBER Working Papers: 18054.

Beaulieu, E. and C. Magee (2004). "Four Simple Tests of Campaign Contributions and Trade Policy Preferences." *Economics and Politics* **16**(2): 163-187.

Bernhofen, D. M. and J. C. Brown (2004). "A Direct Test of the Theory of Comparative Advantage: The Case of Japan." *Journal of Political Economy* **112**(1): 48-67.

Bernhofen, D. M. and J. C. Brown (2005). "An Empirical Assessment of the Comparative Advantage Gains from Trade: Evidence from Japan." *American Economic Review* **95**(1): 208-225.

Brown, D. J. and R. L. Matzkin (1996). "Testable Restrictions on the Equilibrium Manifold." *Econometrica* **64**(6): 1249-1262.

Chō, Y., K. Shōda, et al. (1979). *Meiji chūki sangyō undō shiryō*. Tōkyō, Nihon Keizai Hyōronsha.

de Bavier, E. (1874). *La sériciculture : le commerce des soies et des graines et l'industrie de la soie au Japon*. Lyon and Milan, H. Georg Du Molard frères.

Deardorff, A. V. (1994). Overview of the Stolper-Samuelson Theorem. *The Stolper-Samuelson theorem: A golden jubilee*. A. V. Deardorff and R. M. Stern, With the assistance of Sundari R. Baru. Studies in International Trade Policy. Ann Arbor: University of Michigan Press**:** 7-34.

Deardorff, A. V., R. M. Stern, et al. (1994). *The Stolper-Samuelson theorem : a golden jubilee*. Ann Arbor, University of Michigan.

Debreu, G. (1974). "Excess demand functions." *Journal of Mathematical Economics* **1**(1): 15-21.

Doi, S. (1984). Kinsei tatara seitetsu *Saikō to yakin*. K. e. a. Amakasu. Tokyo, Nihon Hyōronsha**:** 69-103.

Donaldson, D. (2013). "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* **forthcoming**.

Ernst, M. D. (2004). "Permutation Methods: A Basis for Exact Inference." *Statistical Science* **19**(4): 676-685.

Ethier, W. (1974). "Some of the Theorems of International Trade with Many Goods and Factors." *Journal of International Economics* **4**(2): 199-206.

Ethier, W. (1982). "The General Role of Factor Intensity in Theorems of International Trade." *Economics Letters* **10**(3-4): 337-342.

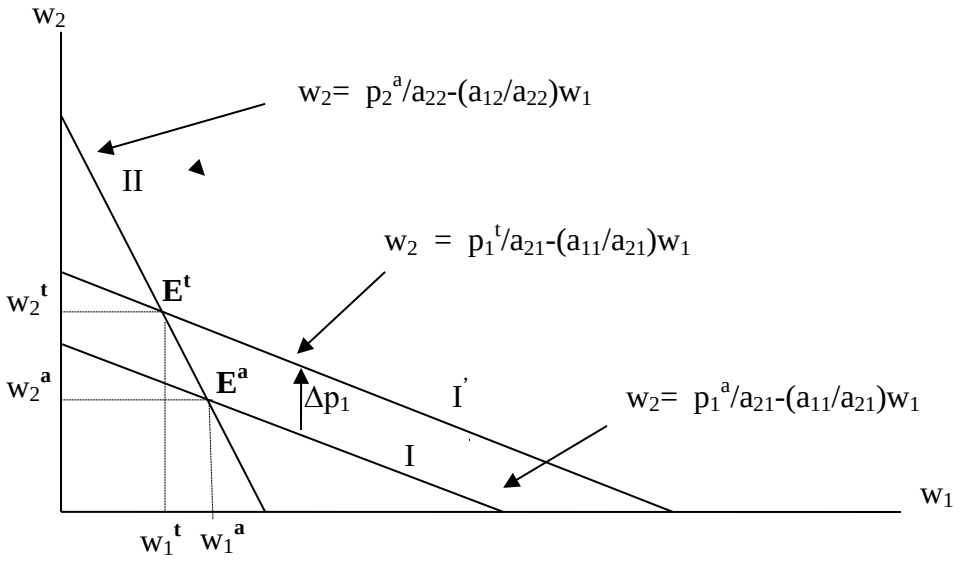
Ethier, W. (1984). Higher Dimensional Issues in Trade Theory. *Handbook of International Economics*. R. W. Jones and P. B. Kenen, Elsvier. **1:** 131-184.

Feenstra, R. C. and G. H. Hanson (1999). "The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979-1990." *Quarterly Journal of Economics* **114**(3): 907-940.
Geerts, A. J. C. (1874). "Useful Minerals and Metallurgy of the Japanese." *Transactions of the Asiatic Society of Japan* **3**(1): 1-16, 27-51,85-97.
German Consul in Edo (1873). Jahresbericht des Konsulats zu Yedo für 1872. *Preussisches Handelsarchiv*. Berlin. **2:** 681-685.
German Consul in Hiogo (1873). Handelsverkehr von Hiogo und Osaka im 1. Halbjahr 1873. *Preussisches Handelsarchiv*. Berlin. **2:** 550-554.
Goldberg, P. K. and N. Pavcnik (2007). "Distributional Effects of Globalization in Developing Countries." *Journal of Economic Literature* **45**(1): 39-82.
Gribble, H. (1874). "The Preparation of Vegetable Wax." *Transactions of the Asiatic Society of Japan* **3**(1): 98-101.
Gunma-ken (1977). *Gunma kenshi. Shiryō hen*. Maebashi-shi Gunma-ken
Heckman, J. J. (2000). "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective." *The Quarterly Journal of Economics* **115**(1): 45-97.
Honda, I., Ed. (1936). *Nihon sanshigyō shi* Tokyo, Dai Nihon Sanshika.
Itoh, M. and M. Tanimoto (1998). Rural Entrepreneurs in the Cotton-Weaving Industry of Japan. *Toward the rural-based development of commerce and industry: Selected experiences from East Asia*. Y. Hayami. Washington, D.C., World Bank**:** 47-68.
Jones, R. W. and J. A. Scheinkman (1977). "The Relevance of the Two-Sector Production Model in Trade Theory." *Journal of Political Economy* **85**(5): 909-935.
Katz, L. F. and D. H. Autor (1999). Changes in the Wage Structure and Earnings Inequality. *Handbook of labor economics. Volume 3A*. O. Ashenfelter and D. Card, Handbooks in Economics, vol. 5. Amsterdam; New York and Oxford:Elsevier Science, North-Holland**:** 1463-1555.
Katz, L. F. and K. M. Murphy (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics* **107**(1): 35-78.
Krugman, P. R. (2000). "Technology, Trade and Factor Prices." *Journal of International Economics* **50**(1): 51-71.
Krugman, P. R. (2008). "Trade and Wages, Reconsidered." *Brookings Papers on Economic Activity*: 103-138.
Le Gendre, C. W. (1878). *Progressive Japan, a study of the political and social needs of the empire*. New York and Yokohama, C. Lévy.
Ludbrook, J. and H. Dudley (1998). "Why Permutation Tests Are Superior to t and F Tests in Biomedical Research." *The American Statistician* **52**(2): 127-132.
Lyman, B. S. (1879). *Geological Survey of Japan: Reports of Progress for 1878 and 1879*. Tookei, Public Works Department.
Magee, S. (1980). Three Simple Tests of the Stolper-Samuelson Theorem. *Essays in Honor of Harry G. Johnson*. P. OPpenheimer. London, Oreil.
Mantel, R. R. (1974). "On the characterization of aggregate excess demand." *Journal of Economic Theory* **7**(3): 348-353.

McCloskey, D. N. (1976). "Does the Past Have Useful Economics?" *Journal of Economic Literature* **14**(2): 434-461.
Mitchener, K. J. and M. Ohnuki (2007). "Capital Market Integration in Japan." *Monetary and Economic Studies* **25**(2): 129-153.
Miyamoto, M. (1980). "Bakumatsu Meiji Shoki Kyoto no Bukka Hendo nitsuite: 1830–79, Osaka tono Hikaku nioite (Price Fluctuation from the End of the Tokugawa Era to the Early Meiji Era: Comparison with Osaka 1830–79)." *Osaka Daigaku Keizaigaku* **30**(2): 159-183.
Miyamoto, M. and Ōsaka Daigaku Kinsei Bukkashi Kenkyūkai (1963). *Kinsei Ōsaka no bukka to rishi*. Tōkyō, Sōbunsha.
Nakamura, K. (1962). *Kaitaiki hōken nōson no kenkyū : Suwa-han Imai-mura*. Tokyo, Sōbunsha.
Narita, S., K. Koyana, et al. (1978). *Kogai kinuburui taisei*. Tōkyō, Kōwa Shuppan.
Nihon Senbai Kōsha (1956). *Shōnō senbai shi*. Tōkyō, Nihon Senbai Kōsha.
Nishikawa, S. (1978). "Productivity, Subsistence, and By-Employment in the Mid-Nineteenth Century Choshu." *Explorations in Economic History* **15**(1): 69-83.
Oka, M. and R. Yamazaki (1983). *Nihon keizaishi : bakuhan taisei no keizai kōzō*. Kyōto, Mineruva Shobō.
Plunkett, F. R. (1875). Report by Mr. Plunkett on the Mines of Japan. *House of Commons Sessional Papers***:** 465-519.
Rauch, J. E. (1999). "Networks versus markets in international trade." *Journal of International Economics* **48**(1): 7-35.
Redding, S. J. and D. M. Sturm (2008). "The Costs of Remoteness: Evidence from German Division and Reunification." *American Economic Review* **98**(5): 1766-1797.
Roe, B. E. and D. R. Just (2009). "Internal and External Validity in Economics Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments, and Field Data." *American Journal of Agricultural Economics* **91**(5): 1266-1271.
Saito, O. (2010). "An Industrious Revolution in an East Asian Market Economy? Tokugawa Japan and Implications for the Great Divergence." *Australian Economic History Review* **50**(3): 240-261.
Saitō, O. (1973). "Movements of Agricultural Wage Rates in Japan." *Shakai-Keizai-Shigaku* **32**(2): 50-69.
Saitō, O. (1998). *Chingin to rōdō to seikatsu suijun : Nihon keizaishi ni okeru 18--20-seiki*. Tōkyō, Iwanami Shoten.
Saitō, O. (2009). "Land, labor and market forces in Tokugawa Japan." *Continuity and Change* **24** (1): 169-196.
Saitō, O. and T. Settsu (2006). Money, credit and Smithian growth in Tokugawa Japan. Tokyo, Hitotsubashi University.
Shinbo, H. (1978). *Kinsei no bukka to keizai hatten : zen kōgyōka shakai e no sūryōteki sekkin*. Tōkyō, Tōyō Keizai Shinpōsha.
Sippel, P. (2006). Technology and Change in Japan's Modern Copper Mining Industry. *Institutional and Technological Change in Japan's Economy: Past and Present*. J. Hunter and C. Storz. New York, Routledge**:** 10-26.

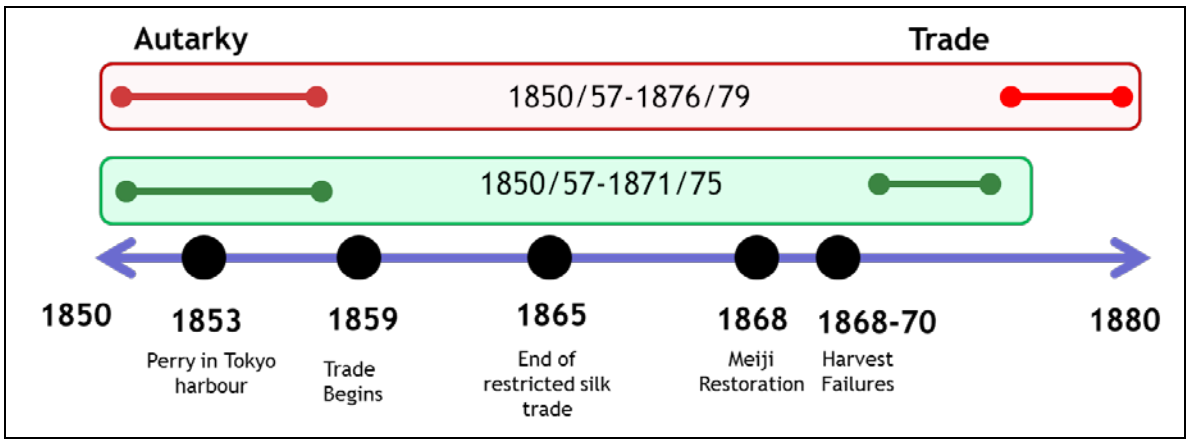
Sonnenschein, H. (1973). "Do Walras' identity and continuity characterize the class of community excess demand functions?" *Journal of Economic Theory* **6**(4): 345-354.
Sonnenschein, H. (1974). "An Axiomatic Characterization of the Price Mechanism." *Econometrica* **42**(3): 425-433.
Stahel, J. (1880). "Cultivation of Cotton and Sugar in Japan." *Commercial Relations of the United States. Reports of the Consuls of the United Stastes*(1): 82.
Stolper, W. F. and P. A. Samuelson (1941). "Protection and real wages." *Review of Economic Studies* **9**: 58-73.
Sugiyama, S. (1988). *Japan's industrialization in the world economy, 1859-1899 : export trade and overseas competition*. London ; Atlantic Highlands, NJ, Athlone Press.
Syrski, S. (1872). Landwirtschaft in Japan. *Fachmännische Berichte über die österreichisch-ungarische Expedition nach Siam, China und Japan (1868-1871)*. K. Scherzer. Stuttgart, Julius Maier**:** 175-227.
Syrski, S. (1872). Über Seidenkultur in Japan. *Fachmännische Berichte über die österreichisch-ungarische Expedition nach Siam, China und Japan (1868-1871)*. K. Scherzer. Stuttgart, Julius Maier**:** 228-286.
Takahashi, I. (1995). Wrought Iron from the Traditonal Tatara Furnace. *History of Industry and Technology of Japan*. E. Pauer**:** 137-152.
Tamura, H. (2001). "The technological development of textile production centers in early Meiji Japan (1870s to 1880s): a case study of the socio-economic impact of the fashion for woollen textiles from western Europe." *Socio-Economic History* **67**(4): 23-48.
Tanimoto, M. (1998). *Nihon ni okeru zairaiteki keizai hatten to orimonogyō : shijō keisei to kazoku keizai*. Nagoya-shi, Nagoya Daigaku Shuppankai.
Tokyo Keizai zasshi. Tokyo, Keizai Zasshisha**.**
Uchida, H. (1988). "Narrow Cotton Stripes and Their Substitutes." *Textile History* **19**(2): 159-170.
Uemura, S. (1986). *Nihon Seitō Gijyutsushi 1700-1900*. Tokyo, Seibundōshuppan.
von Hayek, F. (1964). The Theory of Complex Phenomena. *The Critical Approach to Science and Philosophy*. M. Bunge. London, Collier McMillan**:** 332-349.
Waseda Daigaku Keizaishi Gakkai (1960). *Ashikaga orimono shi*. Ashikaga, Ashikaga Senʾi Dōgyōkai.

Figure 1: The Stolper-Samuelson prediction



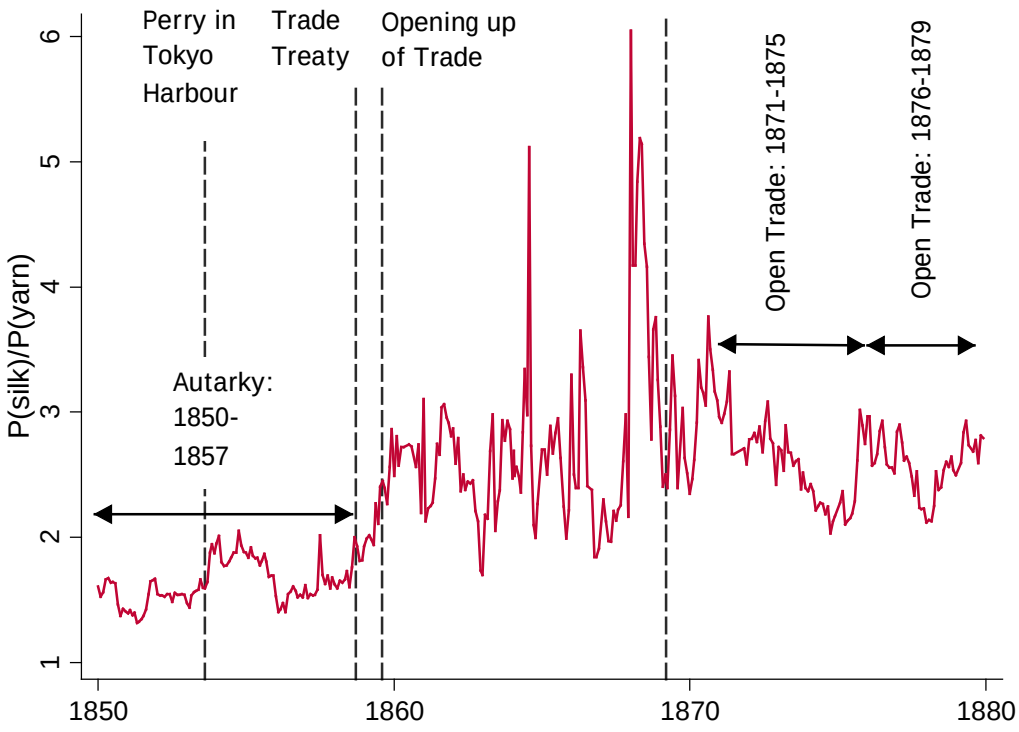
Notes: For a discussion this figure, please see that text.

Figure 2: Major events and experimental windows



Notes: The figure shows the two test intervals: 1850-57 to 1871-75 and 1850-57 to 1876-79. For a discussion of the figure, please see the text.

Figure 3: Relative price shocks from 1850 to 1880: raw silk and cotton yarn

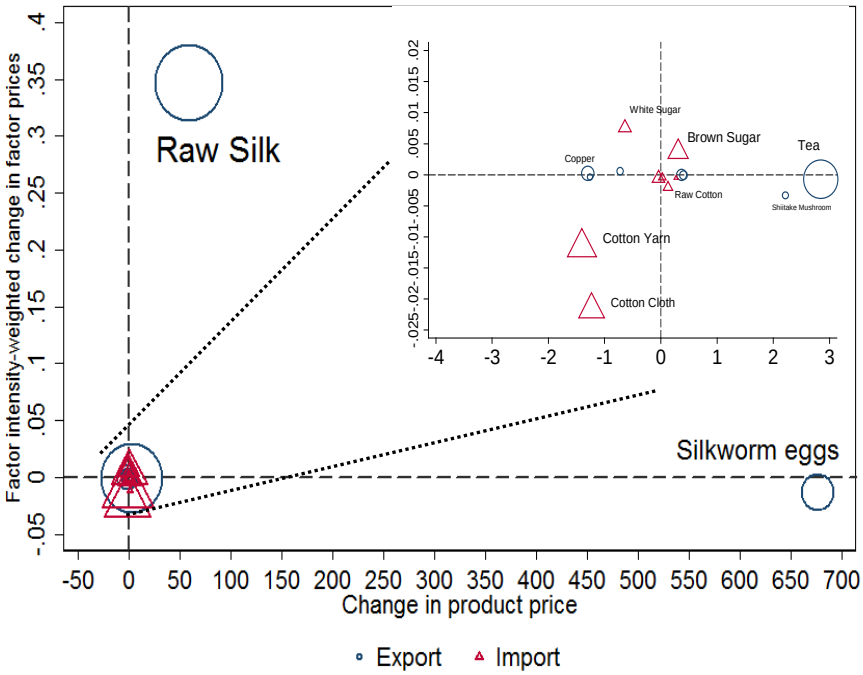


Source: Miyamoto and Ōsaka Daigaku Kinsei Bukkashi Kenkyūkai (1963, Tables 18 and 19)

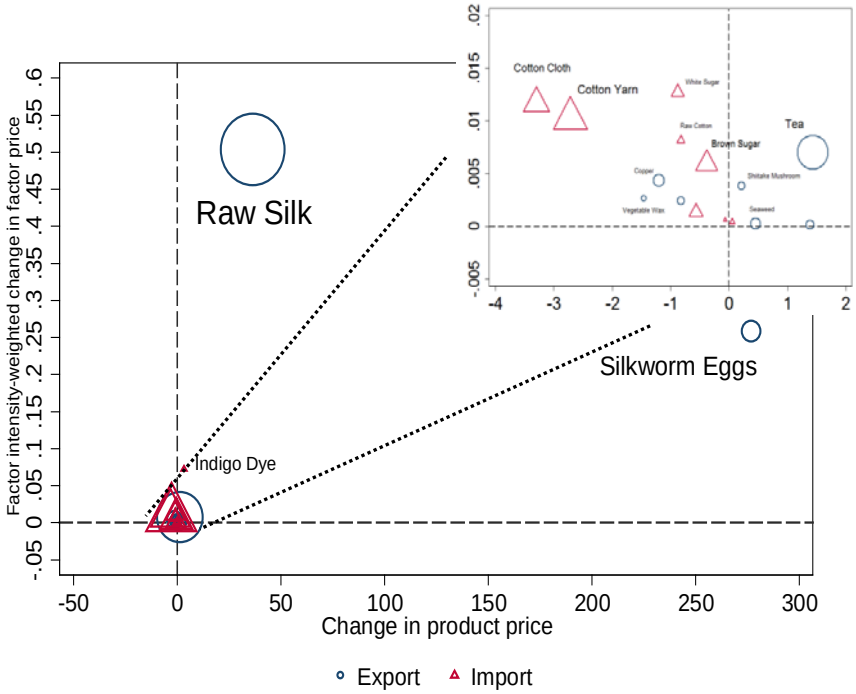
Notes: The prices of yarn (in silver *monme*) are converted to gold ryō using the monthly exchange rates found in Miyamoto and Ōsaka Daigaku Kinsei Bukkashi Kenkyūkai (1963, Table 7)

Figure 4: The pattern of factor and product price changes assuming regionally undifferentiated land

Panel A: 1850-1857 to 1871-1875



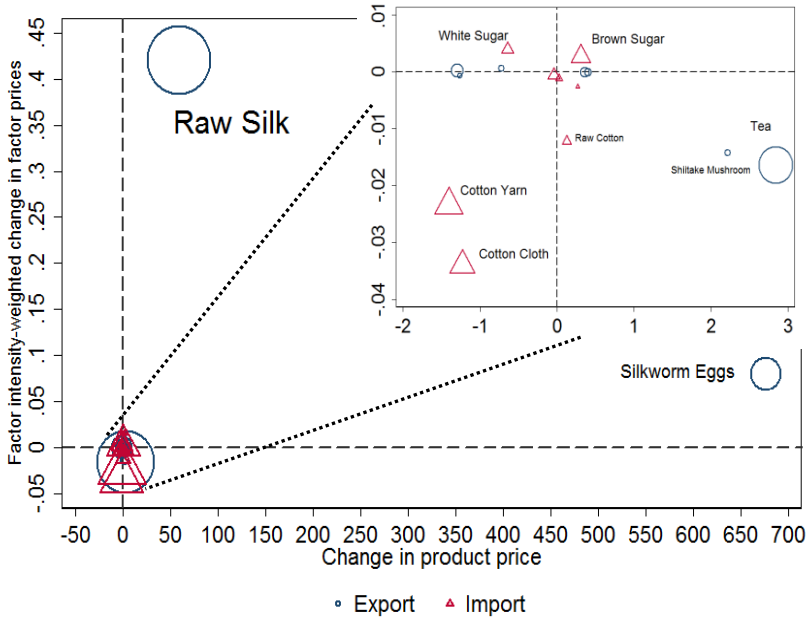
Panel B: 1850-1857 to 1876-1879



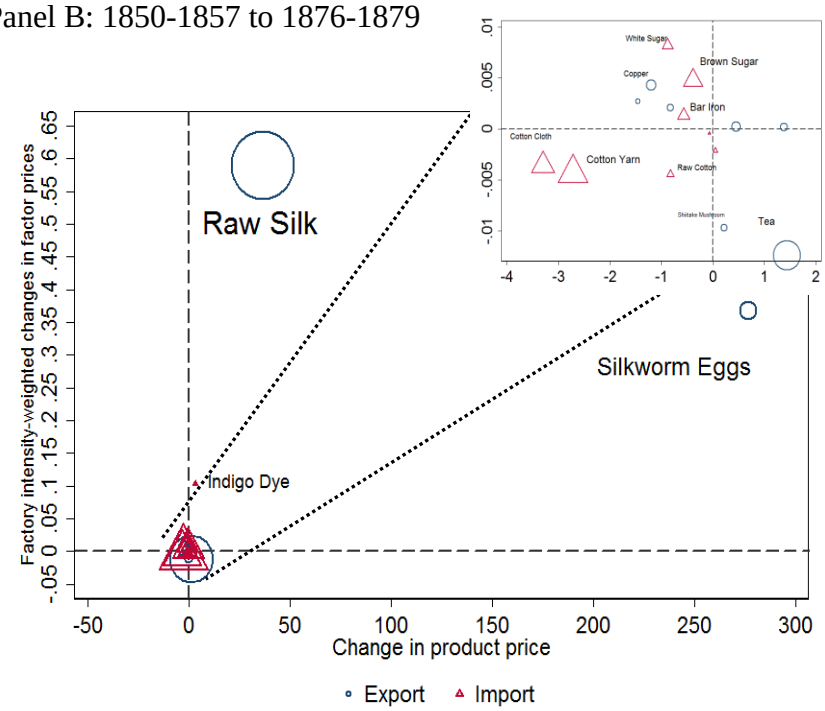
Source: Results of calculations presented in Table 4, Panel A.
Notes: Circles represent the values for exports and triangles represent the values for imports. The size of each marker is proportional to product's share in total trade.

Figure 5: The pattern of factor and product price changes assuming regionally differentiated land

Panel A: 1850-1857 to 1871-1875



Panel B: 1850-1857 to 1876-1879



Source: Results of calculations presented in Table 4, Panel B.
Notes: Circles represent the values for exports and triangles represent the values for imports. The size of each marker is proportional to product's share in total trade.

Table 1: Product scope, Price changes and Sources

Product	Price (in real en per kin)	Δp (in real en/rin per kin) (%Δp in parentheses)		Importance for tradable goods sector (1871-1879)		Frequency of Price Data and source
Imports	1850-1857	1850-57 to 1871-75	1850-57 to 1876-79	Share of M	M as Share of Q+M	
Raw cotton	58.7	1.3 (-8.3)	-8.3 (-14.1)	1.3	0.05	monthly Osaka
Cotton yarn (10s Japanese)	131.7	-14.0 (-10.7)	-27.2 (-20.6)	16.8	0.29	monthly Osaka
Cotton cloth	187.6	-48.0 (-25.6)	-72.4 (-38.6)	11.6	0.34	annual Osaka
Brown sugar	15.5	3.1 (19.9)	-3.8 (-24.8)	7.5	0.70	monthly Osaka
White sugar	28.0	-6.4 (-22.8)	-8.8 (-31.3)	2.8		annual Tokyo
Rice	5.9	0.3 (4.3)	-0.7 (-11.3)	0.6	0.01	monthly Osaka
Soy beans and other legumes	8.8	2.7 (30.9)	0.5 (5.7)	0.4	0.005	monthly Osaka
Indigo dye	51.8	9.0 (17.3)	32.3 (62.3)	0.04		annual Osaka
Bar Iron	24.4	-0.4 (-1.6)	-5.6 (-23.1)	2.9	0.70	monthly Osaka
Exports				Share of X	Share of Q	
Raw silk	1333.5	593.5 (44.5)	366.5 (27.5)	35.9	0.49	monthly Osaka
Silkworm eggs	4622.0	6768.0 (146.4)	2768.0 (59.9)	5.7	0.44	annual Merchants and seasonal consular†
Tea	54.7	28.4	14.4	24.3	0.94	annual Osaka

		(52.0)	(26.3)			
Copper	63.5	-13.0	-12.0	3.2	0.40	annual Osaka
		(-20.4)	(-18.9)			
Camphor	46.3	-12.6	-8.2	0.9		annual Nagasaki sales and seasonal consular†
		(-27.2)	(-17.8)			
Vegetable wax	44.0	-7.2	-14.5	0.9		annual Osaka
		(-16.5)	(-33.1)			
Shiitake mushroom	81.1	22.2	2.2	1.0		annual Osaka
		(27.3)	(2.7)			
Cut seaweed	4.1	3.7	4.5	2.2		annual consular
		(88.8)	(109.5)			
Cuttlefish	30.0	4.0	30.0	1.5		annual consular
		(13.4)	(46.2)			

†Seasonal observations numbering about 5-6 per year.

Notes: All prices changes are in real *en* of 1855-1856 per kin. One kin is equal to 1.33 British lbs. The en was worth one-thousandth of a gold ryō, which was equivalent to one yen when it replace by the yen in 1871. Rice prices per *koku* (a measure of volume) were converted to kin using the data from Atkinson (1881) and soy prices per *koku* were converted using data from the United States Department of Agriculture. Cloth prices for one *tan* of cloth were converted to the equivalent in kin by assuming that one *tan* of cloth used 95 *monme* of yarn (or about 0.8 lbs.).

Sources: Miyamoto and Ōsaka Daigaku Kinsei Bukkashi Kenkyūkai (1963) provides monthly or annual notations of goods from the Osaka market that include the period 1850-1879. Camphor prices during autarky are the purchase prices of the bakufu (the Shogunate) at Nagasaki for 1850 and 1851/1852 found in Nihon Senbai Kōsha (1956).Prices for the 1870s are average market prices reported in German and British consular reports. The cut seaweed price during autarky is for June, 1859 just at the beginning of open trade in Kanagawa as reported by the British consul. The autarky prices during the 1870s are average export prices at the treaty ports and market prices provide by the German and British consuls. The records of the average sales price of a silkworm egg merchant found in Honda (1936, pp. 105-107) were used to estimate prices in autarky. Subsequent prices are for eggs from the production regions of Oshiu, Sinshiu and Joshiu reported in de Bavier (1874), Syrski (1872) and the British and German consular reports for the 1870s.

Stahel (1880) for cotton and sugar production, Sugiyama (1988, Tables 4-1, 4-2 and 5-2) for raw silk and tea production, Geerts (1874, p. 16) for iron production, Itoh and Tanimoto (1998) for cotton cloth production and consumption and Le Gendre (1878, Appendix Table) for production of rice and beans and Sippel (2006, Table 2.1).

Table 2: **Tests of the Stolper-Samuelson Theorem for Two Test Periods with Regionally Differentiated and Undifferentiated land**

	1850/57-1871/75		1850/57-1876/79	
Factor	$\sum_j \boldsymbol{a_{ij}\Delta p_j}$	Δw_i	$\sum_j \boldsymbol{a_{ij}\Delta p_j}$	Δw_i
Panel A: Regionally undifferentiated land				
Skilled Male (day)	32.1	-0.002	13.4	0.001
Unskilled Male (day)	169.5	-0.002	71.5	0.008
Female (day)	9.1	-0.003	3.8	0.0002
Capital	15.0	0.062	7.0	0.054
Land (tan)	2.4	-0.130	1.0	-0.034
Δw(AΔp)		**0.119**		**0.900**
Panel B: Regionally Differentiated land				
Skilled Male	32.1	-0.002	13.4	0.001
Unskilled Male	169.5	-0.002	71.5	0.008
Female	9.1	-0.003	3.8	0.000
Capital	15.0	0.062	7.0	0.054
Land in the west (Kansai)	-0.0001	-0.860	-0.001	-0.940
Land in the east (Kantō)	2.4	0.146	1.8	0.290
Δw(AΔp)		**0.790**		**1.459**

Notes: **Δw(AΔp)** is the test value for the economy expressed in ryō/yen. The land in the east is used as a specific factor for the production of products of the sericulture industry and indigo. The land in the west is used for all other products except for rice. Rice is produced using land from both the east and the west.

Source: For a discussion of the calculations, please see the text.

Table 3: Testing for the Robustness of the Stolper-Samuelson Results with Permutation Tests

Assumption on Land	Results of permutation tests	1850-1857 to 1871-1875		1850-1857 to 1876-1879	
		With Rent excluded from permutations	With Rent included in permutations	With Rent excluded from permutations	With Rent included in permutations
Regionally undifferentiated land	Type I error (**α**)	0.083	0.185	0.001	0.045
	N permutations	5,985	5,985	1,001	1,001
Regionally differentiated land	Type I error (**α**)	0.016	0.067	0.002	0.048
	N permutations	5,985	5,985	1,001	1,001

Source: Results of Monte Carlo sampling from permutations and calculations of permutations.
Notes: The limited number of observations on female wages constrained the number of permutations to 5,985 for 1871-1875 and to 1,001 in 1876-1879.

Appendix 1: Using permutation tests to assess Type I error

Permutation test methods offer a non-parametric approach to use sample data to test a variety of hypotheses, including the null hypothesis about the equality of test statistics indicated by equation (6).[31] The approach used in this paper makes only one assumption about the data used in the test: the available prices and wages available from the historical record are randomly drawn from the vector of goods and factor prices prevailing in equilibrium in period 1 and in period 2 (or 3). As the discussion that follows suggests, this assumption allows for results that are free of distributional assumptions and allow for complete control of the relevant errors so that the Type I error can be viewed as an exact error.

The null hypothesis in this paper is that the difference in the inner products calculated from averages of period-specific vectors of prices and wages should not be significantly affected by whether the prices are drawn from the first or second period. Or alternatively, assigning the prices to the "before" and "after" periods and then calculating the test statistic could very well yield values that are *greater* than the values observed and presented in Table 4. The total number of such possible arrangements is of course the permutation of the count of goods or factor prices observed for the period 1 and the count observed for period 2 (or $\binom{n_1 + n_2}{n_1}$ possible permutations). Each of these permutations

[31] Ludbrook and Dudley (1998) and Ernst (2004) provide some historical background on the use of permutation tests. We view our implementation as a "population test," which assumes that the data that does become available to to us from the past are randomly drawn from the population of prices and wages. Ernst (2004) notes that conditional upon the sample, the test is exact without any assumptions on the distributions that generated the sample observations.

occurs with a probability of $\frac{1}{\frac{(n_1+n_2)!}{n_1!n_2!}}$. For a simple example, suppose that for period 1 historical sources have provided exactly 10 observations on the rent of land in eastern Japan (1850-57) and 9 observations for period 2 (1871-1875). Holding all other factor prices and the prices of raw silk and silkworm eggs constant, there are 92,378 possible ways in which these 19 observations on rent can be assigned to vectors of 10 for period 1 and 9 for period two. The null hypothesis implies that factor prices—including rent— do not differ between the two periods (in a predictable manner). Holding all other prices constant, an exact approach to a permutation test would calculate the value of the inner product based upon the average value of rent for each one of the *purely random* vectors assigned to each period. The relevant α for the test (the probability of Type I error) is the percent of the inner products calculated under the (null) hypothesis of randomness that exceeds the value of the test statistic appearing in Table 4.

The results reported in Table 6 extend this basic approach in two directions. First, following the logic outlined in Ernst (2004, p. 682), Monte Carlo sampling of potential permutations (10,000 samples) was carried out in most cases.[32] In addition, the reported tests involve up to eight possible different groups of permutations (for two product prices and six factor prices). The appropriate permutations were generated for each price variable and then randomly grouped together in order to calculate the test statistic.

[32] The test remains exact and conditionally distribution-free (see Ernst (2004)).