INTER-COUNTRY SPILLOVERS IN THE DIFFUSION OF NEW TECHNOLOGY – AN EPIDEMIC MODEL OF STEAM AND MOTOR SHIP DIFFUSION

By

Anni-Maria Pulkki[1]
Warwick Business School
University of Warwick
Coventry
CV4 7AL

Anni.Pulkki05@phd.wbs.ac.uk
http://go.warwick.ac.uk/ep/pg/bsreba

This version 28 March 2008

---

[1] With grateful thanks to Professor Paul Stoneman for his supervision and the Economic and Social Research Council for financial assistance. All errors are the sole responsibility of the author.

ABSTRACT

We propose that the diffusion of a new technology within a country is influenced by the usage of that technology elsewhere. We extend the classic epidemic diffusion model of Bass (1969) by considering a world diffusion effect, with world diffusion measured as a simple average across countries. The hypothesis is tested using country-level annual data on steam- and motor ship diffusion in 15 countries. We find support for the world diffusion effect however the theory does not indicate the sign of the effect, and indeed we find both positive and negative estimates. Unfortunately we cannot control for market segmentation (i.e. the relative importance of long- vs. short-haul traffic) with the present data. However, we can conclude that the relationship between domestic and world diffusion does not conform to a simple leader-laggard model. We argue that difficulties in interpreting the results may also be due to the nonstationarity of the variables, a feature consistently ignored in the epidemic literature. We conclude that the epidemic framework is an appropriate starting point for the testing of the world diffusion effect but that the hypothesis should also be tested using an alternative theoretical model of diffusion.

Keywords: Technology diffusion, epidemic model, spillovers, steamships

CONTENTS

1 Introduction

In the epidemic approach the use of a new technology increases as information about its existence and characteristics becomes more widespread. There are two seminal pieces of work in the literature. Bass (1969) argued that potential adopters learn about the existence and characteristics of a new technology from those who have already adopted it. When the number of users is small, opportunities to learn about the technology are few, but as the number of users increases the flow of information also increases. According to Bass, it is this spreading of information that drives diffusion. His model has been extensively used in the marketing literature that analyses the diffusion of consumer goods. The second seminal work is Mansfield (1961) who developed a model where the uncertainty of adoption discourages diffusion. As more information about the technology is obtained, uncertainty about the profitability of the new technology is reduced, thus increasing its diffusion. Mansfield's work has been influential in the economics literature. In this paper, we refer to Bass and Mansfield's models as the classic models of diffusion.

We analyse the diffusion of steam and motor ships over the period 1809-1939 during which sailing ships were gradually eliminated from world commercial shipping. The diffusion of steamships is our main concern since motor ships did not begin their diffusion until the early 20th century. Unfortunately for most of the 15 countries in our dataset we do not have data on steam and motor ship tonnages separately.[2] Therefore we consider these two together as the set of new technologies that superseded sailing ships. When the time period permits us to assume that the data does not include motor ships, we will refer exclusively to steamships in our discussion.

Steam engines are commonly identified in the literature as a General Purpose Technology that had wide-ranging effects on productivity (e.g. Breshanan and Trajtenberg 1995) however surprisingly the literature on steamship diffusion is rather limited. In particular, there is a nearly complete lack of empirical studies that model the diffusion process. The exceptions are Cohn (2005) and Comin, et al. (2006). Comin, Hobijn and Rovito (2006) are interested in classifying a large number of technology-country pairs according to the fit of the logistic curve. Using the same data as we here, they fit a logistic curve to two measures: the

---

[2] A separate analysis of steamship diffusion would have been possible only for Belgium, Denmark and Norway. Motor ships are separated from steamships also in figures for Finland, Germany, and the United Kingdom but this separation is incomplete in the early years.

proportion of steam and motor ships in total tonnage, and the steam and motor ship tonnage per capita. The authors do not report the logistic parameter estimates for individual technologies but point out that the logistic does not fit the latter measure well because of a "moving ceiling" (Comin, et al. 2006:18). Cohn (2005) regresses the yearly change in gross steamship tonnage on the lagged volume of immigration from Europe to the United States. His simple linear regression model is constructed without reference to epidemic or other theoretical models of diffusion. Cohn finds support for a positive relationship between immigration and change in gross tonnage, which he interprets as a response by shipbuilders to demand, and also a positive effect from a time dummy in 1870 which he argues reflects the opening of the Suez Canal and the diffusion of the compound engine. Most studies on steamship diffusion have followed Graham (1956) and Harley (1971) and focused on the gradual technological change in both steam and sailing ships and changes in the relative prices (typically freight rates) without attempting to model the diffusion process empirically.

The main technological problem with early steamships was low fuel efficiency. The space taken up by coal in the hull was considerable especially on long journeys, and the price of coal at refuelling stations increased with the distance to Britain which was the main source of coal at the time (Harley 1971). Steamships were initially able to compete with sailing ships only on short journeys and in the transport of passengers and perishable or valuable cargo for which speed and reliability commanded a premium price. The soiling of the bottom of the ship slowed down early steamships at a time when sailing ships increased their speed with the help of Maury's current and wind charts published in 1850 (Graham 1956). Soon however speed and reliability became the main advantages of the steamship. Sailing ships were most competitive in the transport of bulk cargo on distant journeys such as from Europe to India, China, Australia and the west coast of America. Even the opening of the Suez Canal in 1869 did not mark a turning point for the Age of Sail because of the low fuel efficiency of steamships. This was despite the fact that sailing ships were unable to use the Canal and instead travelled around the Cape of Good Hope. For British shipping, the Suez Canal had the biggest effect on China trade which was dominated by tea, a high value good, whereas a quarter of the Bengal and Burma trade by tonnage was still sail in the early 1890s (Harley 1971).

The main technological improvements in steamship technology were the diffusion of high-pressure engines and developments in metallurgy. The perfection of the high pressure design was reached in stages with gradual fuel efficiency improving at each stage. Steel structures became available from the 1870s which enable the use of boiler plates and steel tubes that could withstand yet higher pressures. As the price of steel fell to rival that of best quality iron this led to the development of the triple expansion engine in the 1880s (Graham 1956). Over the period of technological improvements the ship size increased and the cost of ships declined. Crew size could be reduced because steel ships were more reliable, required less crew to operate, and were easier to navigate. Altogether these mainly incremental improvements increased the distance margin at which cargo steamers and sailing ships competed on equal terms (Harley 1988).

Benefits of sailing ships included their faster turnaround time and ability to exploit trade winds. Many of the improvements that took place in steamship technology also occurred in sailing ships. For example, there was an increase in the size and reliability of ships due to the use of iron and steel. Wooden ships were enhanced with metal structures and later iron and steel sailing ships were built. In harbours, the movement of these very large ships was made possible by steam tugs. Capacity per crewman was still higher in sailing ships than steamships in the late 1880s, although the gap was narrowing (Knauerhase 1968). Labour productivity increased due to economies of scale from bigger ships and the higher reliability of steel masts and rigging (Harley 1988). In sailing ships, the use of auxiliary steam engines also increased labour productivity. There is a strand of literature that considers these technological improvements in sailing ship technology to have been an irrational response by sailing ship owners to the threat of the new superior steamship technology. However, Howells (2002) has argued that the historical facts do not support this “sailing ship effect” as it has later become known. He argues that the exploitation of iron and steel could as well have been a result of competition among rival sailing ship companies as an “irrational” reaction to the threat of steam.

We can attempt to distinguish two alternative views regarding why the diffusion of the steamship took as long as it did. The first view is purported by most studies in economic history, for example Graham (1956) and Harley (1971, 1988) and Crafts (2004). According to this view, incremental innovation in steamship technology was crucial for steamship diffusion. As their productivity improved, steamships replaced sailing ships gradually according to the relative cost of shipping by steam and sail. A lower cost was preferred for most cargo. It follows that steamship usage at any given point in time in a particular market was in equilibrium. The view is supported by analyses suggesting that both sailing ship and steamship markets were highly competitive (Broeze 1975, Harley 1971).

An alternative approach regards diffusion as a disequilibrium adjustment process. A view in line with epidemic theory would argue that it was the spreading of information and knowledge about the superiority of steam (and motor) ships that was driving the end of the Age of Sail. Some early steamships had sails (Howells 2002), which may indicate uncertainty about the benefits of steam; although it can also be seen as a rational response to the unreliability or high costs of early steamships. If investors were uninformed about the true profitability of steamships and ship owners faced capital constraints, the market would not have been in equilibrium. This appears to have been an issue in steamship diffusion in Norway for example, but not in the United Kingdom (Harley 1971 discussion, Samstag and Joshi 2005). There is also evidence of suspicion or even prejudice towards steamships by contemporaries; for example, Lloyds[3] viewed steamships with "marked distrust" (Graham 1956:74)**.** We will base our extension of the classic epidemic models on this second type of reasoning, although the first approach will be used in order to interpret the results that we obtain.

Before we proceed it is worth making a note about the deed of nationality of commercial vessels. National shipping, meaning nationally owned and manned commercial vessels, dominated during the 19th and early 20th centuries (Barton 1999). Beginning in the inter-war period but particularly since the Second World War shipping companies have sought to reduce their costs by pursuing shipping registries, so-called "flags of convenience". This has marked the end of national merchant shipping. Examples of convenience flagging for political and military reasons can be found far back in history however it was in the interwar period that the first US and European ships were flagged to Panama for purely economic reasons (Alderton and Winchester 2002). While it would be inappropriate to take countries as the units of analysis for today's shipping industry our analysis of steam and motor ship diffusion concerns the period before 1939 and so we consider it appropriate to do so.

The structure of the paper is the following. First, we construct two measures of steam- and motor ship diffusion: one within-country and one world diffusion measure. Second, we discuss the diffusion of steam- and motor ships and the data used in this study. Third, we

---

[3] The Lloyds' Register in London

present our extended model of within-country diffusion which incorporates international spillovers. The empirical analysis consists of the model's time series properties and the regression results.

2 Measuring diffusion

Diffusion studies use a variety of different measures depending on the nature of the particular technology, such as product or process, as well as the objectives of the researcher and the data available to them. In economic history, tonnage figures are most commonly used to analyse the shipping industry. We have data on the net registered capacity, which is an estimate obtained by subtracting from the total enclosed space (i.e. gross capacity) a portion devoted to engines, crew's quarters etc. The net registered capacity is a British method of measurement that was introduced in the mid-1850s and subsequently imitated around the world. One ton is equal to 100 cubic feet or 2.83 cubic meters. Had we wanted to analyse steam and motor ships as a product technology, we could have used data on the number of registered ships. However this data is problematic for example because the average steamship capacity changed considerably during the period. We regard steam and motor ships as a process rather than product technology, that is, as a method of production, and therefore a measure of the cargo-carrying capacity is justified.

Having chosen to use tonnage data we face a second choice between a level measure and a proportional measure. A level measure of diffusion is for example the tonnage of steam and motor ships in a given country whereas a proportional measure is the share of steam and motor ships in the country's total tonnage. In our dataset a country's total use of merchant ships consists of sailing, steam and motor ships. It follows that the proportional measure of steam and motor ship diffusion is in fact a measure of the "disappearance" of sailing ships which we are interested in. If there were more alternative technologies a level measure may have been a better choice.

Let T(i,t) denote the total tonnage of commercial ships and let S(i,t) denote the tonnage of commercial steam and motor ships in country i at time t. The proportion of steam and motor ships

$$\frac{S(i,t)}{T(i,t)}$$

is our measure of diffusion. Given the periods we study, S(i,t)/T(i,t) tends to unity over time as the number of sailing ships diminishes towards zero. Note that a similar approach could be applied to other technologies. Our case is an example of a process technology where T(i,t) is the total production or capital use at time t and S(i,t) is the proportion embodying steam or motor ship technology.

We now introduce the concept of an asymptotic level of diffusion. It is common in epidemic studies to treat the asymptotic level as a constant to which the level of usage converges as diffusion progresses over time. This amounts to arguing that usage is in disequilibrium during the process of diffusion. An alternative approach that we prefer is to allow the asymptotic level to be time-dependent. We can think of the asymptotic level increasing as there are changes to each of the different ship technologies and as a variety of external factors varies the size of the country's commercial fleet. We denote the asymptotic level by

$$\frac{S^*(i,t)}{T^*(i,t)}.$$

This definition allows the asymptotic level to vary both across countries and across time. The significance of the asymptotic level in epidemic models is that the discrepancy between the actual and the asymptotic diffusion drives diffusion forward, as we will see below.

We now suggest a way of measuring international diffusion that can be used to test the hypothesis that intra-country diffusion is influenced by diffusion elsewhere. Let k be the number of countries in our sample; we will refer to these countries as the "world". Let W(t) denote the sum of steam and motor ship tonnages at time t in these k countries, and let TW(t) denote the total commercial tonnage in the same countries. World diffusion of steam and motor ships is the ratio of W(t) to TW(t), i.e. the simple average of the level of diffusion in the k countries at time t:

$$\frac{W(t)}{TW(t)} = \frac{\sum_{i=1}^{k} S(i,t)}{\sum_{i=1}^{k} T(i,t)}.$$

In our data W(t)/TW(t) tends to unity over time as sailing ships disappear from the world commercial fleet. In order to test the spillover hypothesis we need a measure of use elsewhere, that is, a measure of world diffusion that is not directly affected by use in country i. We use simply

$$\frac{\sum_{j=1, j\neq i}^{k} S(j,t)}{\sum_{j=1, j\neq i}^{k} T(j,t)} = \frac{W(t) - S(i,t)}{TW(t) - T(i,t)}.$$

Let there be an asymptotic world steam and motor tonnage W*(t) and an asymptotic total tonnage TW*(t) which we define as simple averages of the asymptotic levels in the k countries in period t. A ratio of these asymptotic levels,

$$\frac{W*(t)}{TW*(t)} = \frac{\sum_{i=1}^{k} S*(i,t)}{\sum_{i=1}^{k} T*(i,t)},$$

is the asymptotic level of world diffusion. This varies over time as both W*(t) and TW(*(t) vary.

We have chosen not to weight the contributions of individual countries in the world diffusion measure and therefore W(t)/TW(t) is dominated by countries with a large tonnage. That is, even dramatic changes in diffusion in small countries will not greatly influence world diffusion. Whether or not this property is desirable depends on our theoretical view of how world diffusion affects domestic diffusion. The unweighed measure is consistent with an epidemic model where information about steam and motor ships, the new technology, is primarily obtained from countries with the biggest commercial fleet. Following the reasoning of Bass and Mansfield, the sheer number of adopters matters, therefore adoption by a small country will not give much information whereas a high level of diffusion in a large country is a strong signal that the new technology is profitable. A more sophisticated version would consider some sub-group of the k countries that is thought to be more influential than the others. Literature on spillover effects could be used to argue that diffusion in countries with close trade relations or short geographical distance is more likely to influence domestic diffusion. A weighted measure would require a substantial theoretical argument about the nature of information flows between countries, possibly a country-by-country analysis of the relevant factors, and certainly the development of a novel weighting scheme since such does not exist in the literature. We merely attempt a first test of the hypothesis that world diffusion affects domestic diffusion and consider that an unweighted measure is appropriate for this purpose. We will however discuss the contribution of individual countries to this measure in some detail (see section 3.2).

3 The diffusion of steamship technology

3.1 Steam and motor ship data

Our data is from HCCTAD (Historical Cross-Country Technology Adoption Dataset) which has been compiled by Diego Comin and Bart Hobijn using data provided by Mitchell (1998).[4,5] The countries analysed are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Netherlands, New Zealand, Norway, Sweden, United Kingdom, and United States. The study period is 1809-1939. These dates are chosen by us with the objective that the whole diffusion process is covered from beginning to end in all countries; however data availability presents some constraints to this at both ends of the period. Let us first discuss the start date. Table 1 tabulates the first observations of steam and motor tonnage and total tonnage for each of the 15 countries. Total tonnage is typically available at an earlier date than steamship tonnage however missing observations are frequently a problem in this early data. Because consecutive annual data is available from the year of the first steam and motor tonnage (with the exception of Canada and Finland)[6] we set this as the start of the study period for each country. It follows that differences in the start date partly reflect real differences in the beginning of diffusion and partly the fact that data is not available at low tonnage levels. Table 1 reveals that with the exception of Canada the dataset captures the beginning of the diffusion process very successfully. This is important because estimation of epidemic models is sensitive to any early missing observations.

Although the data is of good quality there was a need for some imputations. We did this by randomly choosing a tonnage level between the values just before and after the missing year.[7] All years for which more than two consecutive values were imputed for a given country are reported in Table 1.

---

[4] HCCTAD includes data on Japan but we could not use this because total tonnage and steam and motor tonnage are measured differently to each other. More countries are available in: Brian R. Mitchell, *International Historical Statistics*, 4 ed. (Basingstoke: Palgrave Macmillan, 1998). However, the quality of the additional data is insufficient for our purposes.

[5] HCCTAD is freely available at http://www.nber.org/hccta. It was first used by Comin and Hobijn in: Diego Comin and Bart Hobijn, "Cross-Country Technology Adoption: Making the Theories Face the Facts," *Journal of Monetary Economics* 51 (2004).

[6] For Canada and Finland the start dates are 1892 and 1873 respectively.

[7] Where several consecutive figures were missing, random values were assigned to the missing years in ascending order over time for steam and motor ships and in descending order for sail ships.

Table 1. First observations and imputed years

Country	*Steam and motor*		*Total tonnage*		*Imputations*
	first year	*tons*	*first year*	*tons*	*(years)*
Australia	1876	20	1871	97	
Austria	1837	1	1829	188	
Belgium	1837	1	1837	23	1914-1919
Canada	1867*	46	1867	768	
Denmark	1844	1	1829	60	
Finland	1848*	1	1842	120	1888-1891
France	1838	10	1837	697	
Germany	1850	4	1829	265	1914-1922
Italy	1862	10	1862	654	1901-6, 1918-34
Netherlands	1846	2	1846	380	
New Zealand	1870	6	1857	7	
Norway	1866	6	1830	135	
Sweden	1865	12	1830	131	1866-1869
United Kingdom	1815	1	1788	1278	
United States	1809	1	1789	202	

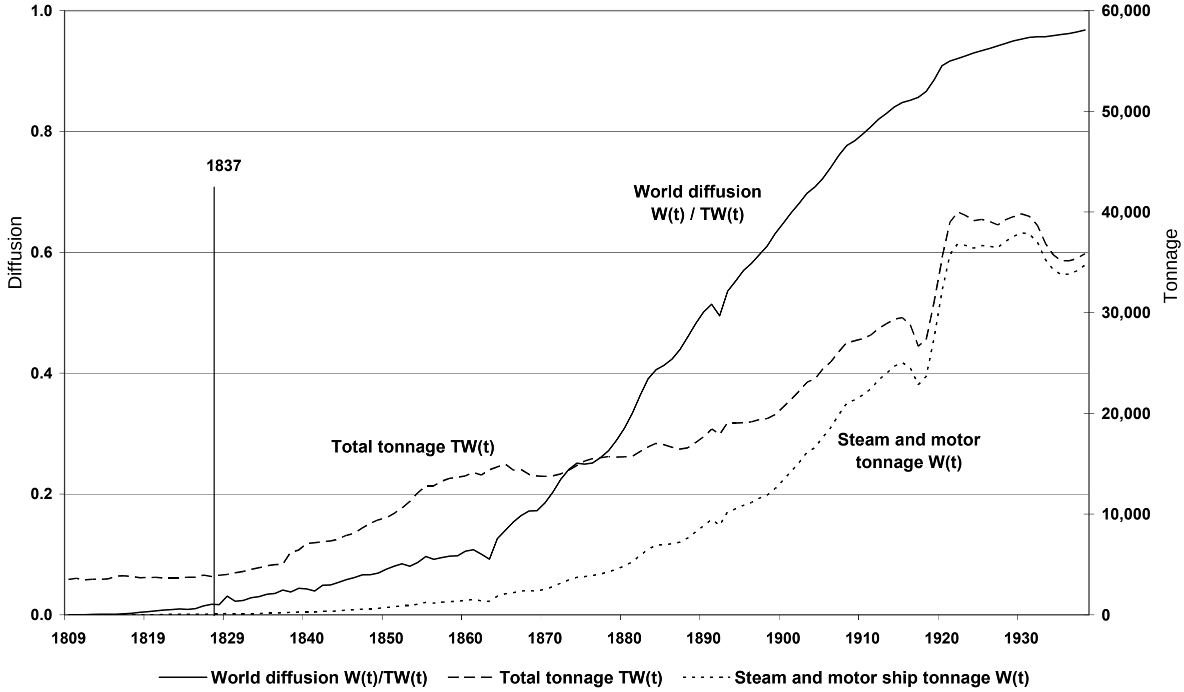
Notes to Table 1: The last column indicates dates for which more than two consecutive values were imputed. *Canadian and Finnish data are available only occasionally until 1892 and 1873 respectively.

Turning to the world diffusion measure W(t)/TW(t), we chose 1837 as the start date for the study period. This date is an arbitrary choice based on the availability of data on total tonnage (see Table 1). For those countries where no data was available until after 1837 we either imputed a small number of early values for total tonnage or assumed that tonnage was negligible in order to construct W(t)/TW(t). This was straightforward except for Canada and Italy where the first figures for total tonnage are for 1867 and 1862 respectively. We do not have enough information to impute earlier values and neither can we assume that tonnage levels were non-negligible in 1837. For these reasons we decide to exclude Canada and Italy from the world diffusion measure. This means that W(t)/TW(t) is defined as the average of diffusion in 13 countries (i.e. k=13).[8] We have computed this measure for earlier years starting in 1809 but have not made any imputations. It follows that there are jumps in TW(t) in the years 1829 and 1837 when data on new countries becomes available. Nevertheless the early data can be used for illustrative purposes. The tonnage levels W(t) and TW(t) and world diffusion W(t)/TW(t) are plotted for the period 1809-1938 in Figure 1.

---

[8] The other 13 countries can be included in the measure of world diffusion for one of the following reasons: 1) all tonnage data is available in 1837 or very close to that date thus it can be imputed; 2) sail tonnage is available and steam and motor tonnage can be inferred to have been minimal; 3) both are missing yet the first available data suggests that missing values are small in magnitude. We have sail ship tonnage for 7 countries in 1837 (Austria, Belgium, Denmark, France, Sweden, United Kingdom, United States) and impute a value for this year for 4 more countries (Finland, Germany, Netherlands, Norway). Two countries are inferred to have had minimal tonnages (Australia, New Zealand).

Figure 1. World diffusion of steam and motor ships (1809-1938)



World diffusion of steam and motor ships shows the typical S-shaped diffusion curve: it increases slowly at first then at a higher rate from about the mid-1860s before slowing down again as diffusion approaches the maximum of unity. This pattern is consistent with the argument that steamships were at the peak of their importance between 1870 and the First World War.[9] Note that even in 1938 there were some sailing ships in the world commercial fleet.[10] The shape of the diffusion curve suggests that fitting a logistic curve, which is the basis of classic epidemic models, may be appropriate. Figure 1 also supports the choice of 1837 as the start date because W(t)/TW(t) and W(t) are at a very low level: the tonnage of steamships was 237 tons which was merely 3.8 per cent of the total tonnage. The apparent increase in the mid-1860s in the speed of diffusion coincides with the diffusion of the high-pressure steam engine. From this period until the First World War total world tonnage

---

[9] Referring to the contribution of steamships to British economic growth, this argument has been made by Nicholas Crafts, "Steam as a General Purpose Technology: A Growth Accounting Perspective," *The Economic Journal* 114 (2004).

[10] The sail ship tonnage exceeded 400 tons in Canada and the United Kingdom, 200 tons in the United States and 100 tons in France and Germany. In Britain sailing ships carried small perishable cargoes using small harbours until the 1930s when the use of large ports and lorries for inland transport finally drove sailing ships out of these trades Basil Greenhill, *The Merchant Schooners* (Newton Abbot: David & Charles, 1968)..

increases but steam and motor tonnage increases at a greater rate, implying a seemingly steady increase in diffusion.[11]

During the First World War tonnage ships were destroyed in large numbers in many countries which is apparent as a drop in tonnages in Figure 1. After the war, there is a period of sharp increase which is also reflected in W(t)/TW(t), although the deviation from the long-term trend in the latter is not as apparent. Some commentators have argued that the First World War hastened the end of the sailing ships (e.g. Samstag and Joshi 2005). Our raw data does not support this claim. However, if the causal factors behind diffusion were different during the war, we need to be concerned. One indication of this possibility is the increase in motor ship tonnage while sail and steam tonnage were falling.[12] Plots of S(i,t)/T(i,t) also reveal discontinuities however these are not of the same type in all countries (e.g. a plateau is observed in Germany and Sweden). We conclude that the arguments for excluding the First World War years are insufficient and therefore we extend the study period to also cover these years. However we also estimate all models over a shorter period ending in 1913 with the exception of Canada for which the study period would have become too short. A comparison of the estimates from these two periods will be used to judge whether the short period is more reliable.

We have chosen to end the long study period at the start of the Second World War because diffusion was close to its maximum at this time and so inclusion of subsequent years would not have brought more information. Diffusion had reached a level of at least 90 per cent in all countries except Canada by 1939 (Figure 2). Indeed, four countries had already reached this level in 1913 which suggests that the shorter period may be sufficient and the inclusion of the additional years unnecessary. Although 1939 was the first year of the Second World War the data suggests a break in 1940 and so there is no strong reason to exclude 1939 data from the study period. The exceptions are Germany and the Netherlands for which 1939 data is missing, and Norway where 1939 tonnage levels are considerably higher than in earlier years.

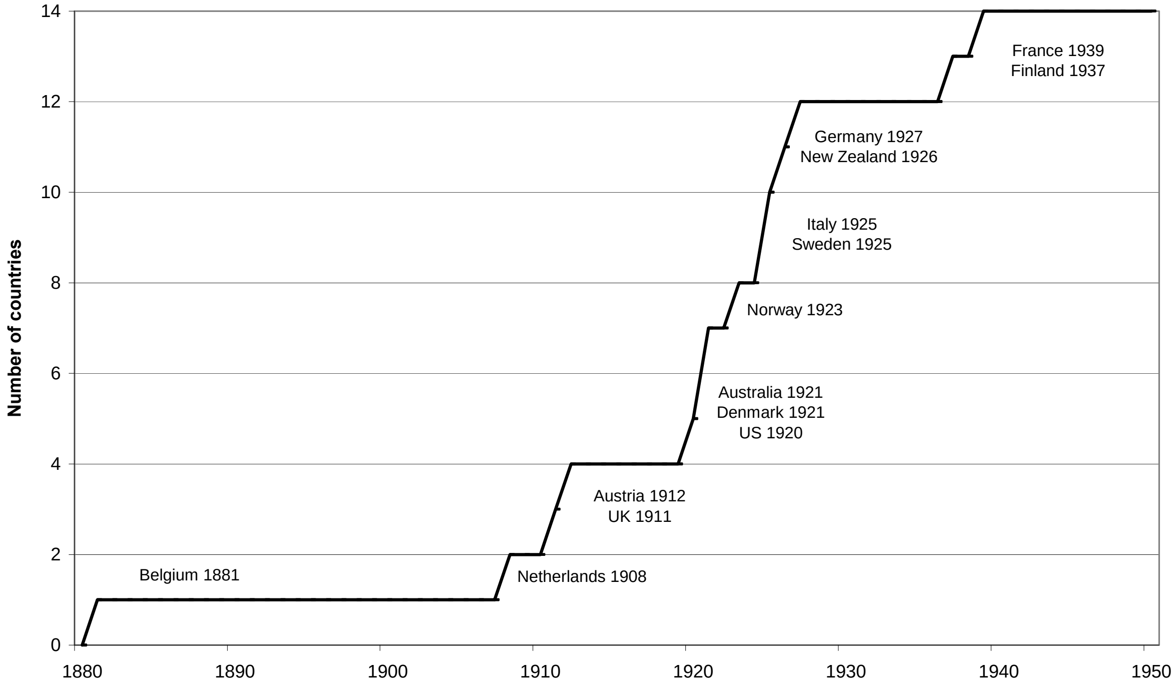
---

[11] The exception is 1892 which is due to a fall of the order of 1000 tons in the United Kingdom. We suspect that this is a mistake but have been unable to confirm it. The same figure appears in Brian R. Mitchell, *International Historical Statistics*, 5 ed. (Basingstoke: Palgrave Macmillan, 2003).

[12] Denmark and Norway are the only countries for which we have data on motor ships during the First World War.

For these reasons the study period ends in 1938 for world diffusion, Germany, Netherlands and Norway, and in 1939 for all other countries.[13]

Figure 2. Year in which 90 per cent diffusion was reached ($S(i,t)/T(i,t) \geq 0.9$)



Notes to Figure 2: Canada is excluded from this figure because the highest level of diffusion during the study period is 0.66 in 1939.

The study periods for each country are given in Table 2. There was no data after 1912 for Austria so only a short period was estimated. The long period ends in 1925 for Italy because there were no observations between 1926-1934; values were imputed but only used in computing $W(t)/TW(t)$. Since our model (see below) assumes that diffusion is increasing throughout the study period, we end the long study period in 1925 for Australia where diffusion fell after this year, and in 1931 for Belgium where tonnage levels fell considerably. However for all other countries we considered two study periods with the end marked by the start of the First and Second World Wars respectively.

---

[13] Although 1938 is the final observation for world diffusion we can take 1939 as the final period for most countries because $W(t)/TW(t)$ enters as a lagged value in the extended model (see below).

Table 2. Study periods

Country	*Start date(s)*	*End date(s)*
Australia	1876	1913, 1925
Austria	1837	1912
Belgium	1837	1913, 1931
Canada	1892	1939
Denmark	1844	1913, 1939
Finland	1873	1913, 1939
France	1838	1913, 1939
Germany	1850	1913, 1938
Italy	1862	1913, 1925
Netherlands	1846	1913, 1938
New Zealand	1870	1913, 1939
Norway	1866	1913, 1938
Sweden	1865	1913, 1939
United Kingdom	1815, 1837	1913, 1939
United States	1809, 1937	1913, 1939

3.2 World diffusion

Due to the way it is constructed, our measure of world diffusion W(t)/TW(t) is dominated by the two countries with the biggest merchant ship tonnage, namely the United Kingdom and the United States. Their combined share was never under 60 percent in either W(t) or TW(t) during the period 1837-1938.[14] The United States initially dominated steamship tonnage contributing over half of the world tonnage until the end of the 1860's. Thereafter the United Kingdom was dominant until the First World War. She had the biggest total as well as steam and motor tonnage in 1870-1919. The share of the United Kingdom in W(t) exceeded 40 per cent in every year during this period except 1919 and 50 per cent in 1874-1903. This was the period in which most of the growth in world steam and motor ship diffusion occurred. A group of the six biggest countries can be identified which constitutes at least 89 per cent of both measures throughout the period 1837-1938. France is initially the third most important country in W(t) until she is overtaken by Germany in 1889, Norway in 1906 and Netherlands in 1921. Figures for the relative importance of the six countries in selected years are given in Table 3.

---

[14] The Norwegian steam and motor ship tonnage increased considerably in 1939 (by 74 percent or over 2000 tonnes) which results in the US and UK shares falling below 60 percent in that year.

Table 3. Six biggest contributors to W(t) and TW(t) (selected years)

Steam and motor tonnage W(t) and total tonnage TW(t) (tons)

Measure	*1840*	*1860*	*1880*	*1900*	*1920*	*1938*
W(t)	303	1,454	4,846	13,370	32,132	34,697
TW(t)	7,071	13,789	15,688	20,637	35,359	35,848

Contribution of selected countries to W(t)

	1840	*1860*	*1880*	*1900*	*1920*	*1938*
France	3.3%	4.7%	5.7%	3.9%	3.4%	4.3%
Germany	0.0%	1.6%	4.0%	9.9%	6.3%	6.8%
Netherlands	0.0%	0.8%	1.3%	2.0%	3.0%	6.4%
Norway	0.0%	0.0%	1.2%	3.8%	4.1%	8.0%
United Kingdom	29.0%	31.2%	56.2%	51.7%	33.5%	29.7%
United States	66.7%	59.7%	25.0%	19.9%	43.0%	34.6%
Total (6 countries)	99.0%	97.9%	93.5%	91.2%	93.4%	89.8%

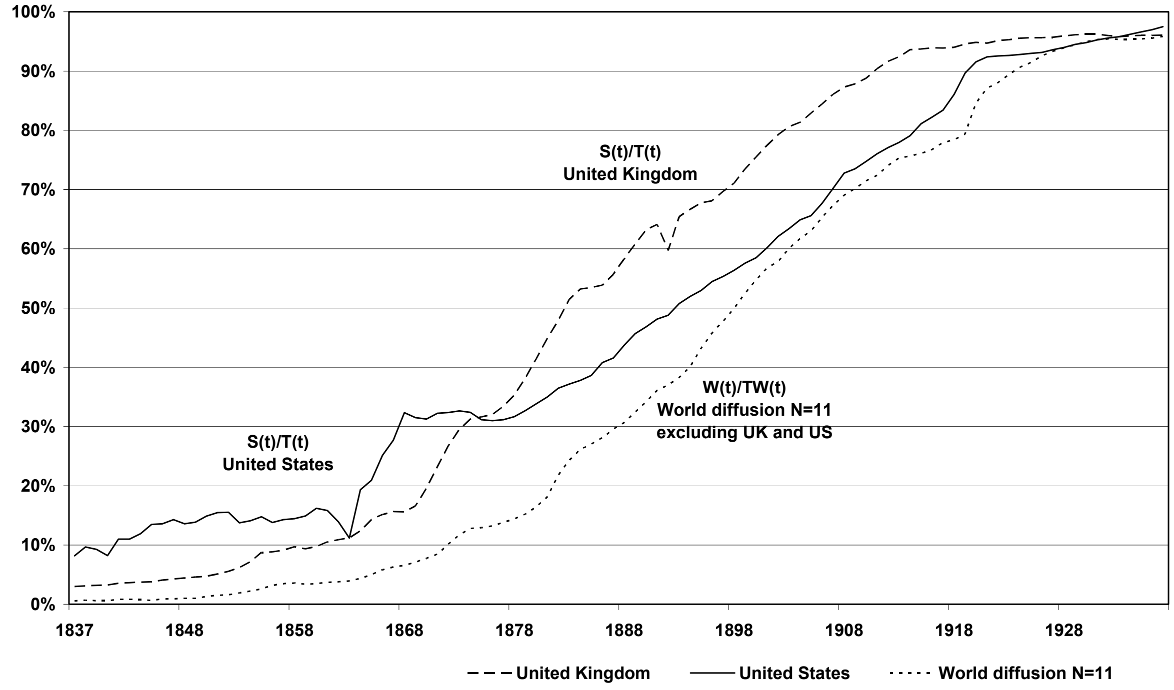
Contribution of selected countries to TW(t)

Country	*1840*	*1860*	*1880*	*1900*	*1920*	*1938*
France	9.4%	7.2%	5.9%	5.0%	4.3%	4.6%
Germany	5.0%	5.6%	7.5%	9.2%	6.6%	6.9%
Netherlands	4.8%	3.6%	2.1%	1.7%	2.8%	6.2%
Norway	2.9%	3.9%	9.7%	7.3%	4.3%	7.8%
United Kingdom	39.1%	33.8%	41.9%	44.4%	32.1%	29.9%
United States	30.8%	38.8%	22.8%	22.0%	42.7%	34.2%
Total (6 countries)	92.0%	92.9%	89.9%	89.7%	92.8%	89.7%

Given the characteristics of the world diffusion measure let us consider the expected relationship between this and the diffusion process in each individual country. The hypothesis of the epidemic framework is that information flows drive diffusion. It is assumed that users of a new technology have more knowledge than non-adopters and that this knowledge is supportive of adoption. Thus we would expect that in countries lagging behind world diffusion knowledge from abroad encourages adoption at home. Conversely, even if domestic diffusion is ahead of world diffusion we would not expect a negative relationship because knowledge and experience accumulated by foreign adopters can only reduce the uncertainty faced by domestic sailing ship owners about the profitability of adoption. We could argue that ship owners in countries ahead of world diffusion will not obtain much addition information from adoption by others, thus the relationship will be weak. One such country is the United Kingdom where diffusion was above world diffusion for most of the period. The United States also had a high level of diffusion and if the United Kingdom were not included in the world diffusion measure the United States would also have been ahead of world diffusion throughout. This is illustrated in Figure 3 where we have plotted S(i,t)/T(i,t) in the United

Kingdom and United States and world diffusion W(t)/TW(t) without the contribution of these two countries.

Figure 3. World diffusion W(t)/TW(t) without the United Kingdom and the United States (1837-1938)



The reasoning above may be consistent with Bass and Mansfield's arguments but the history of shipping suggests a more complicated picture. We have argued above that the substitution of sailing ships by steamships occurred gradually, starting with shorter routes and the transport of passengers and valuable cargo. As steam technology (in particular fuel efficiency) improved steamships were adopted on ever longer routes. This suggests that if information was crucial for diffusion, ship owners would have given more weight to information from adopters on those routes and in those markets that they themselves were operating in. Ideally, we would like to analyse steam and motor ship diffusion in each of the different markets separately, i.e. broken down by the length of the journey and the nature of the cargo. We would still expect world usage to have a positive effect but the relationship is likely not to conform to a simple leader/laggard scenario.

In fact, negative relationships could also occur. Consider the following hypothetical scenario. Suppose that a majority of ship owners in a given country operate on short routes most amenable to early steamships and that they all have converted to steam. The remaining sail ship owners operate on distant journeys carrying bulk cargo. Suppose that most ship owners abroad are also involved in this latter type of trade. Now the high level of domestic diffusion suggests that the remaining ship owners should be adopting steamships at a high rate. However, only information from adopters in the same market is relevant. The low level of world diffusion will discourage further domestic diffusion. In this scenario, we may find that the estimated world diffusion effect is negative since it counters the positive effect of high domestic diffusion. We cannot test the hypothesis suggested by this scenario because we only have aggregate data. However, we do know that in practice diffusion proceeded one market segment at a time, which was due to the long period of incremental technological improvements that made steamships ever more competitive against sailing ships.

4 Epidemic model of diffusion

Applying the Bass model (1969) to our empirical example we model the evolution of steam- and motor ship tonnage S(i,t) by

$$dS(i,t) = \left(\gamma + \beta \frac{S(i,t)}{S^*(i,t)}\right)\left(S^*(i,t) - S(i,t)\right). \tag{1}$$

This says that diffusion increases in every period t in which the asymptotic tonnage S*(i,t) exceeds the actual tonnage S(i,t).[15] The parameter β captures the positive effect on diffusion of information-spreading by those who have already adopted (the "endogenous" speed of diffusion). The second parameter γ is called the "exogenous" effect because Bass interpreted this as the effect on diffusion of an information flow from a (domestic) source other than current users, such as advertising.

The Bass model nests another classic model of intra-country diffusion, the Mansfield model. Mansfield (1961) argued that diffusion is driven by the reduction in uncertainty that results from current (domestic) usage. The Mansfield model is simply the Bass model with the parameter γ equal to zero. In both models if S*(i,t) were constant over time the time path of S(i,t) would follow an S-shape. When S(i,t) is small relative to S*(i,t), diffusion proceeds at an increasing rate until an inflexion point, after which the closer S(i,t) is to S*(i,t) the slower is diffusion.

We propose a general model that incorporates the effect of use elsewhere. The general model is a simple extension of the Bass model, and is given by

---

[15] It is assumed that S(i,t) is always less than S*(i,t), i.e. dS(i,t) is never negative.

$$dS(i,t) = \left(\gamma + \beta \frac{S(i,t)}{S^*(i,t)} + \alpha \frac{W(t) - S(i,t)}{W^*(t) - S^*(i,t)} \right) \left(S^*(i,t) - S(i,t) \right). \tag{2}$$

Setting α=1 we have the Bass model and setting α=γ=0 we have the Mansfield model. Since we are interested in the proportional measure of diffusion, S(i,t)/T(i,t), we write the model as

$$\frac{dS(i,t)}{T(i,t)} = \left(\gamma + \beta \frac{S(i,t)/T(i,t)}{S^*(i,t)/T(i,t)} + \alpha \frac{(W(t) - S(i,t))/(TW(t) - T(i,t)}{(W^*(t) - S^*(i,t))/(TW(t) - T(i,t))} \right) \left(\frac{S^*(i,t) - S(i,t)}{T(i,t)} \right). \tag{3}$$

We assume that the asymptotic steam and motor tonnage in period t is equal to the total tonnage in that period:

$$S^*(i,t) = T(i,t).$$

The reason for this assumption is that we do not have data on S*(i,t). In the literature, S*(i,t) has either been treated as a constant or modelled as a function of other variables. Our assumption says that sailing ships are inferior to steam and motor ships in every period. Similarly, we assume that W*(t)=TW(t) which allows us to write the model as

$$\frac{dS(i,t)}{T(i,t)} = \left(\gamma + \beta \frac{S(i,t)}{T(i,t)} + \alpha \frac{W(t) - S(i,t)}{TW(t) - T(i,t)} \right) \left(1 - \frac{S(i,t)}{T(i,t)} \right). \tag{4}$$

Discretization by the standard Euler approach gives the estimation equation for the general model as

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \left(\gamma + \beta \frac{S_{i,t-1}}{T_{i,t-1}} + \alpha \frac{W_{t-1} - S_{i,t-1}}{TW_{t-1} - T_{i,t-1}} \right) \left(1 - \frac{S_{i,t-1}}{T_{i,t-1}} \right) + \varepsilon_{i,t} \tag{5}$$

where $\varepsilon_{i,t} \sim N(0,\sigma^2)$. We use NLS as the estimation method.

We also estimate a model of world diffusion to serve as a benchmark to which the results for individual countries can be compared. Using the same procedure as above, the Bass model of world diffusion is given by

$$\frac{dW(t)}{TW(t)} = \left(\gamma + \beta \frac{W(t)}{TW(t)} \right) \left(1 - \frac{W(t)}{TW(t)} \right), \tag{6}$$

for which the discrete version is

$$\frac{W_t - W_{t-1}}{TW_{t-1}} = \left(\gamma + \beta \frac{W_{t-1}}{TW_{t-1}} \right) \left(1 - \frac{W_{t-1}}{TW_{t-1}} \right) + \varepsilon_t \tag{7}$$

where $\varepsilon_t \sim N(0,\sigma^2)$. The Mansfield model of world diffusion is simply a restricted version of the Bass model with γ=0.

We have referred above to two classic models of intra-country diffusion, the Bass and the Mansfield models. There is no international effect here, since intra-country diffusion is affected only by domestic usage. Both models and their variants have been extensively applied in diffusion studies at various levels of aggregation. In general, the country-level has not been as popular as lower levels of aggregation such as the regional, industry or firm-level. For an example of a country-level study see Gruber and Verboven (2001); for the use of epidemic models in the economics and marketing literatures see Geroski (2000) and for the sociological literature see Strang and Soule (1998). We apply the model not only to the country level but also to the international level; that is, we also fit the Bass and Mansfield models to our data on world diffusion of steam and motor ships.

The extended model we have proposed nests the Bass and Mansfield models as special cases. In Bass' terminology, we have incorporated the world effect as an "exogenous learning effect", that is, it affects diffusion in a similar way to the parameter γ. Thus we propose that information from abroad influences the domestic diffusion process. As potential adopters travel, see foreign advertisements and read newspapers they obtain information crucial for adoption. In today's information society such a proposition is hardly contestable but we propose that historically in the context of shipping ship owners probably had reasonable access to information from around the world, whether first- or second-hand. We have chosen not to let spillovers affect the endogenous diffusion speed β. Kumar and Krishnan (2002) argued that diffusion in one country is likely to affect another country's diffusion through both of the parameters γ and β; the latter would mean that "the degree to which a potential adopter would place faith on the internally generated information will be affected by what happens in other countries" (Kumar and Krishnan 2002:321). In our example, we interpret this as that sailing ship owners would treat information from domestic steamship owners with suspicion until world steamship diffusion is at a sufficiently convincing level. This does not seem convincing to us and therefore we extend the classic epidemic models using a simple additive term as proposed above.

For the two countries excluded from our world diffusion measure, Canada and Italy, a measure of diffusion elsewhere is simply the original world diffusion measure, $W(t)$. Indeed the difference $W(t)-S(i,t)$ would have a different meaning from that intended in (4). The estimating equation for the general model for Canada and Italy is then given by

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \left(\gamma + \beta \frac{S_{i,t-1}}{T_{i,t-1}} + \alpha \frac{W_{t-1}}{TW_{t-1}} \right) \left(1 - \frac{S_{i,t-1}}{T_{i,t-1}} \right) + \varepsilon_{i,t} \tag{8}$$

where $\varepsilon_{i,t} \sim N(0,\sigma^2)$. With $\gamma=0$ this gives the extended Mansfield model for Canada and Italy.

To test the explanatory power of the classic epidemic models we consider also a theoretical model where diffusion is driven exclusively by world usage. This doesn't have a corollary in the literature. Setting $\gamma=\beta=0$ in the general model we have

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \left(\alpha \frac{W_{t-1} - S_{i,t-1}}{TW_{t-1} - T_{i,t-1}} \right) \left(1 - \frac{S_{i,t-1}}{T_{i,t-1}} \right) + \varepsilon_{i,t} \tag{9}$$

where $\varepsilon_{i,t} \sim N(0,\sigma^2)$. We call this the fully coupled model. For Canada and Italy the fully coupled model is given by

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \left(\alpha \frac{W_{t-1}}{TW_{t-1}} \right) \left(1 - \frac{S_{i,t-1}}{T_{i,t-1}} \right) + \varepsilon_{i,t} \,. \tag{10}$$

5 Empirical analysis

5.1 Time-series properties

In accordance with current good practice in empirical economics we begin our data analysis by exploring the time-series properties of the model variables. The stationarity of variables is a concern because of the spurious regression problem. The concept of stationarity and the implications of nonstationarity are discussed further in the Appendix, where we also present the full results of the unit root tests. For each of the 15 countries there are five variables that need to be tested for stationarity: these can be identified by multiplying out (5): [16]

$$\frac{S_{i,t} - S_{i,t-1}}{T_{i,t-1}} = \gamma + (\beta - \gamma)\frac{S_{i,t-1}}{T_{i,t-1}} - \beta \frac{S_{i,t-1}^2}{T_{i,t-1}^2} + \alpha \frac{W_{t-1} - S_{i,t-1}}{TW_{t-1} - T_{i,t-1}} - \alpha \frac{S_{i,t-1}\left(W_{t-1} - S_{i,t-1}\right)}{T_{i,t-1}\left(TW_{t-1} - T_{i,t-1}\right)} + \varepsilon_{i,t} \,. \tag{11}$$

Since the general model nests all the other models the only additional variables that need to be tested are the world diffusion measures. We write out equation (7):

$$\frac{W_t - W_{t-1}}{TW_{t-1}} = \gamma + (\beta - \gamma)\frac{W_{t-1}}{TW_{t-1}} - \beta \frac{W_{t-1}^2}{TW_{t-1}^2} + \varepsilon_{i,t} \,. \tag{12}$$

The three variables here were also tested for stationarity. All tests were conducted for the study periods indicated in Table 2.

---

[16] In the cases of Canada and Italy the third and fourth variables on the right-hand side are W(t-1)/TW(t-1) and S(i,t-1)/T(i,t-1) * W(t-1)/TW(t-1) (see equation (8) above).

First we ran the augmented Dickey-Fuller (ADF) test with the null hypothesis or a unit root and the alternative of stationarity. The power of the test was very low: the null was rarely rejected. This was not too surprising given the length of our study periods and the fact that diffusion data typically has a nonlinear time trend, because the ADF test is known to have low power if the time series has a nonlinear trend or structural breaks. In addition to the ADF tests we chose to run Perron's (1989) dummy variable test which allows for a single structural break. Identifying the date of the structural break T_B was straightforward for Finland, where there was a change in measurement practices in 1919.[17] For other countries we used the plot of S(i,t)/T(i,t) to identify a suggested T_B and we also looked for a structural break around the two World War periods. We then conducted F-tests to determine T_B for each country. In the econometric literature, tests have been developed to deal with nonlinear trends explicitly; we intend to explore these in the future, but have not done so yet.

We find that in the Bass and Mansfield models the left-hand side variable is stationary in most countries but non-stationarity cannot be rejected in the right-hand side variables with a few exceptions. Of the two additional variables that enter the general model (see (11) above) nonstationarity cannot be rejected in the first variable and it is only rejected in the second variable for a few countries. See the Appendix for the full results (Tables A1 and A2).

Our results suggest that any model that nests within it either of the classic models will also suffer from nonstationarity. Alternatives to the logistic function have been considered in the epidemic literature and this could provide one way forward. We have investigated the stationarity of the logistic transformation of S(i,t)/T(i,t) but found that nonstationarity cannot be rejected even in this transformed variable. Thus to derive an alternative epidemic model that would not suffer from nonstationarity appears to be far from straightforward. As our aim in this paper is to test an extended model that is firmly within the epidemic modelling framework we proceed with the extended model that nests the classic models. However, we keep in mind the lack of stationarity when interpreting the results.

5.2 Results

The Bass and Mansfield models fitted to the world diffusion measure W(t)/TW(t) provide a benchmark for the evaluation of results for individual countries. The results are presented in

---

[17] Ships below a certain tonnage were excluded from the statistics from 1919 onwards. This dramatically decreased the number of sail ships since this was the most common type among small vessels.

Table 4 for the study periods 1837-1913 and 1837-1938. The estimates of the "endogenous" diffusion speed parameter β are positive and highly significant. In the Mansfield model for the short period which is our preferred estimate as will be shortly explained, the 95 per cent confidence interval for the magnitude of β is 0.09 to 0.13. The Bass "exogenous effect" parameter γ is not significant. The estimates are very close to zero and also negative contrary to expectations. Our results suggest that dropping γ leads to only a very small reduction in the fit of the model, indicated by a fall in the maximised log likelihood (log L) and an increase in the Residual Sum of Squares (RSS) going from the Bass to the Mansfield model. This supports the conclusion that the Mansfield model is superior to the Bass model. This is a general finding across countries and we will therefore focus our discussion on the Mansfield model.

Table 4. World diffusion estimates of the Bass and Mansfield models

Model	Period	N	β	se(β)	t(β)	γ	se(γ)	t(γ)	RSS	log L
B	1837-1913	76	0.13	0.01	9.0	0.00	0.00	-1.5	0.01	216.7
B	1837-1938	101	0.15	0.03	5.6	-0.01	0.01	-1.2	0.08	219.8
M	1837-1913	76	0.11	0.01	11.6				0.02	215.6
M	1837-1938	101	0.12	0.02	6.8				0.08	219.0

Notes to Table 4: Model B is the Bass model and Model M is the Mansfield model. N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood.

The magnitude of the β estimates from the two study periods is nearly equal. Lower standard errors and higher t-values suggest that the models fit better to the short period. A plot of W(t)/TW(t) in Figure 1 above may suggest an explanation: there is marked discontinuity in W(t) and TW(t) during the First World War and the interwar period but this is not as apparent in W(t)/TW(t). The estimation results suggest that the classic models do not fit the data as well if the study period covers these years. This may suggest that world steam- and motor ship diffusion during 1914-1938 was subject to causal influences other than those considered by our models. We conclude that the short period estimates of the Mansfield model are the appropriate choice as benchmark estimates.

We now turn to discuss the results for individual countries beginning with the general model. Across all countries and the two time periods, we find that the estimates of the Bass coefficient γ are generally close to zero and insignificant at 5 per cent. The estimates are significant only in a few cases: Austria, Belgium (both periods), Finland (short period) and the United States (both periods). These estimates are presented in Table 5. The estimated foreign effect α is positive and significant at 5 per cent for Finland and the United States. The estimate of β is however negative for Finland and the United States, which is contrary to expectations. The estimated foreign effect is also positive for Belgium. The Belgian estimates are problematic because the model fits very poorly: only γ is significant at 5 per cent and RSS is very high. A probable reason for the poor fit for Belgium is that the diffusion process was very fast; the level of 90 per cent was reached already in 1881 (see Figure 2). The Austrian results are problematic because the magnitude of the β estimate is exceptionally large. The foreign effect, though significant, is negative. Finally, the Bass coefficient γ is negative for Belgium and Finland which is also not expected. The results of the general model are not very useful because of the insignificance of γ in most countries, and the negative sign of γ for Belgium and Finland.

Table 5. Country diffusion estimates of the general model with a significant γ

Country	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	γ	se(γ)	t(γ)	RSS	log L
Austria	1837-1912	75	1.06	0.16	6.8	-0.91	0.16	-5.6	0.03	0.01	3.7	0.04	172.1
Belgium	1837-1931	94	0.27	0.22	1.2	0.91	0.51	1.8	-0.09	0.04	-2.4	0.73	94.9
Belgium	1837-1913	76	0.29	0.24	1.2	0.83	0.57	1.5	-0.08	0.04	-2.0	0.67	72.1
Finland	1873-1913	40	-0.38	0.19	-2.0	0.13	0.06	2.2	-0.02	0.01	-1.9	0.00	135.3
United States	1810-1913	103	-0.06	0.04	-1.4	0.14	0.04	3.8	0.01	0.00	1.9	0.02	287.1
United States	1837-1913	76	-0.19	0.08	-2.4	0.23	0.06	3.8	0.02	0.01	2.5	0.02	203.9

Notes to Table 5: N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood.

The extended Mansfield model does considerably better than the general model. We now discuss these results together with those of the Mansfield model with no foreign effect and the fully coupled model. For each country we only present the results for the study period that provided a better fit. For most countries this was the short period; the reader is referred to the Notes to Tables 8, 9 and 10 for the detail. We will not discuss the classic Bass model because the estimates of γ were insignificantly different from zero in all cases except one, Austria.[18] Furthermore, all estimates of γ were negative which is contrary to expectation.

In Table 6 are gathered the results for countries in which the estimated effect of world diffusion is positive. These are: Australia, Finland, France, New Zealand and the United States. The parameter α is significant at 5 per cent in France and the United States only but with wide confidence intervals. Recall that in the general model, α is also significant for Finland. In those countries, namely Finland and the United States, for which the parameter γ is significant, the estimate of α is higher in the general model than in the extended Mansfield

---

[18] The estimate was -0.010 with a standard error of 0.006.

model. Since γ was positive in one case and negative in the other, there is little that we can conclude from this however. Turning to β, the estimates are positive and significant in the Mansfield model but negative and insignificant in the extended Mansfield model, in all cases. The fit of the models indicated by RSS and log L is very similar in both models. The change of the sign of β without an improvement in fit suggests that the effect of spillovers is rather more complicated than what we perhaps expected. A comparison with the fully coupled model reveals that this model fits the data equally well as the other two models: the t-statistics, RSS and log L values are all very similar. The countries in Table 6 are the only cases in which the fully coupled model fits the data slightly better than the Mansfield model. All models fit poorly to the data for New Zealand. The likely reason is that because the total tonnage was initially very small the first observed value of S(i,t)/T(i,t) is relatively high, at 22.5 per cent. There was also a decline in diffusion during the First World War which contradicts the model assumptions; however the model does not fit better to the short period.

Table 6. Country diffusion with a positive estimated world diffusion effect

Country	Model	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	RSS	log L
Australia	M	1876-1925	49	0.16	0.05	3.3				0.23	61.4
Australia	E	1876-1925	49	-0.34	0.57	-0.6	0.43	0.50	0.9	0.23	61.8
Australia	F	1876-1925	49				0.14	0.04	3.3	0.23	61.6
Finland	M	1873-1913	40	0.06	0.01	4.3				0.00	132.2
Finland	E	1873-1913	40	-0.05	0.07	-0.7	0.02	0.01	1.6	0.00	133.5
Finland	F	1873-1913	40				0.01	0.00	4.7	0.00	133.2
France	M	1838-1913	75	0.07	0.01	5.7				0.03	188.0
France	E	1838-1913	75	-0.04	0.06	-0.7	0.10	0.05	2.1	0.03	190.2
France	F	1838-1913	75				0.07	0.01	6.2	0.03	189.9
New Zealand	M	1870-1939	69	0.17	0.07	2.6				0.64	63.3
New Zealand	E	1870-1939	69	-0.64	0.56	-1.2	0.73	0.50	1.5	0.62	64.4
New Zealand	F	1870-1939	69				0.16	0.06	2.8	0.64	63.7
United States	M	1837-1913	76	0.08	0.01	7.1				0.03	193.5
United States	E	1837-1913	76	-0.01	0.03	-0.4	0.11	0.04	3.0	0.02	200.9
United States	F	1837-1913	76				0.09	0.01	8.2	0.02	200.9

Notes to Table 6: Model M is the Mansfield model, E is the extended Mansfield model, and F is the fully coupled model. N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood. France and the United States: the models fit better to the short period (higher log L and t-values, better diagnostic test results[19]). Australia and New Zealand: long period with more data is preferred over short period with equally good (poor) fit. Finland: long period fits poorly because of a change in measurement after 1919.

We now turn to the countries where world diffusion appears to have a negative effect. These are: Austria, Germany, Italy, Netherlands and Sweden. Table 7 presents the results. First we

---

[19] In this case, tests for autocorrelation and autoregressive conditional heteroskedasticity indicated problems in the long period but not in the short period.

can point out that the model fits relatively well: all estimates of β are significant at 5 per cent and there are no high RSS values. The foreign effect α is significant at 5 per cent in Austria, Germany and Sweden and at 10 per cent in Italy. The point estimates of β are large in the extended model compared to the Mansfield model and also compared to the benchmark estimate of 0.11. The Austrian estimate of 0.61 is particularly large with a large standard error, although the estimate is lower than in the extended Bass (Table 5). In contrast, recall that in countries where the world diffusion effect is positive (Table 6), the estimate of β is higher in the Mansfield model (and negative in the extended model). In the fully coupled model the estimated magnitude of the single parameter is smaller than β in the Mansfield model. The maximised log likelihood suggests that the fully coupled model does not fit the data as well as the Mansfield model.

Table 7. Country diffusion with a negative estimated world diffusion effect

Country	Model	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	RSS	log L
Austria	M	1837-1912	75	0.16	0.02	6.5				0.07	156.9
Austria	E	1837-1912	75	0.61	0.11	5.8	-0.40	0.09	-4.4	0.05	165.6
Austria	F	1837-1912	75				0.12	0.02	5.2	0.08	151.6
Germany	M	1850-1913	63	0.20	0.02	12.2				0.03	155.8
Germany	E	1850-1913	63	0.30	0.05	6.3	-0.08	0.04	-2.2	0.02	158.1
Germany	F	1850-1913	63				0.14	0.02	8.8	0.04	142.6
Italy	M	1862-1913	51	0.13	0.02	8.9				0.02	133.9
Italy	E*	1862-1913	51	0.20	0.04	4.7	-0.04	0.02	-1.7	0.01	135.2
Italy	F*	1862-1913	51				0.06	0.01	6.6	0.02	125.8
Netherlands	M	1846-1913	67	0.19	0.03	6.7				0.07	134.3
Netherlands	E	1846-1913	67	0.27	0.07	3.9	-0.07	0.06	-1.2	0.07	135.1
Netherlands	F	1846-1913	67				0.13	0.03	5.1	0.09	128.2
Sweden	M	1865-1913	48	0.15	0.02	8.5				0.02	116.2
Sweden	E	1865-1913	48	0.29	0.07	4.2	-0.10	0.04	-2.1	0.02	118.4
Sweden	F	1865-1913	48				0.09	0.01	6.9	0.03	110.6

Notes to Table 7: Model M is the Mansfield model, E is the extended Mansfield model, and F is the fully coupled model. N is the number of observations. RSS is the residual sum of squares. log L is the maximised log likelihood. * indicates the use of equations (22) and (24) in place of E and F. The models fit better to the short than the long period in each case. Netherlands: poor fit to the long period (RSS around 1.0) is probably explained by the lack of change in diffusion between 1921 and 1938. Italy: fit is better for the short period (RSS smaller, log L higher, less diagnostic problems[20]). Germany and Sweden: t-ratios are higher in the short period. There is no data for Austria after 1912.

Finally, there are five countries for which we find no evidence of a world diffusion effect (Table 8). These are: Belgium, Canada, Denmark, Norway and the United Kingdom. The Belgian results are of little interest since the fit of the models is poor; however, we can note

---

[20] In this case, tests for autocorrelation and autoregressive conditional heteroskedasticity indicated problems in the long but not in the short period, for models M and E*.

that α was positive and significant at 10 per cent in the extended Bass model. For the other countries, the estimates of β are close in magnitude to the estimate of the benchmark model, 0.11. The Mansfield and fully coupled models fit equally well with the parameter estimates also being of similar magnitude.

Table 8. Country diffusion with no world diffusion effect

Country	Model	Period	N	β	se(β)	t(β)	α	se(α)	t(α)	RSS	log L
Belgium	M	1837-1931	94	0.43	0.10	4.2				0.78	92.0
Belgium	E	1837-1931	94	0.44	0.21	2.0	-0.02	0.33	-0.1	0.78	92.0
Belgium	F	1837-1931	94				0.58	0.16	3.6	0.81	89.9
Canada	M	1892-1939	47	0.08	0.02	4.9				0.03	106.7
Canada	E*	1892-1939	47	0.08	0.08	1.0	0.00	0.04	0.0	0.03	106.7
Canada	F*	1892-1939	47				0.04	0.01	4.7	0.03	106.2
Denmark	M	1844-1913	69	0.15	0.02	9.4				0.03	169.4
Denmark	E	1844-1913	69	0.17	0.06	2.8	-0.02	0.05	-0.4	0.03	169.5
Denmark	F	1844-1913	69				0.11	0.01	8.5	0.03	165.6
Norway	M	1866-1913	47	0.13	0.01	12.2				0.01	144.7
Norway	E	1866-1913	47	0.14	0.02	6.4	0.00	0.01	-0.2	0.01	144.7
Norway	F	1866-1913	47				0.05	0.01	7.5	0.01	129.6
United Kingdom	M	1837-1913	76	0.13	0.02	7.8				0.04	176.9
United Kingdom	E	1837-1913	76	0.12	0.07	1.7	0.02	0.09	0.3	0.04	179.7
United Kingdom	F	1837-1913	76				0.18	0.02	7.5	0.04	178.3

Notes to Table 8: Model M is the Mansfield model, E is the extended Mansfield model, and F is the fully coupled model. * indicates the use of equations (22) and (24) in place of E and F. N is the number of observations. RSS is the residual sum of squares and log L is the maximised log likelihood. Belgium: the models fit poorly to both periods so the longer period is presented. Canada: only a long period is estimated. Denmark: models fit poorly to the long period (RSS 0.63, log L around 103) possibly because of a decline in diffusion during the First World War. Norway and United Kingdom: fit is poorer for the long period (log L and t-values are smaller).

To conclude, the models generally fit better to the shorter period that finishes around 1913. The evidence supports the existence of spillovers but we can say little about the sign, since there is an equal number of positive, negative and zero estimates. The sign of the parameter β is different in the extended and Mansfield models. This may suggest that the extended model does not successfully separate out the domestic and foreign effects as we had hoped. The estimate of β in the benchmark model is of a magnitude that we would expect. This reflects the relatively smooth nature of the data series that closely resembles the logistic curve. Most of the other estimates of β in the Mansfield model are of a similar magnitude, though less precise and generally somewhat larger. The length of the study period does not appear to be related to the magnitude, nor is the choice of start or end period, except that the estimates for the United Kingdom and the United States are close to those of the benchmark. This is unsurprising given that data series for these countries are also very long and they are the biggest contributors to W(t)/TW(t) (see table 3). There are three countries to which all models fit poorly: Australia, New Zealand and Belgium. The first two data series begin with a relatively higher first observation of S(i,t)/T(i,t) at over 20 per cent. The logistic curve is known to fit poorly if early data is missing, which may be the explanation here. The Belgian estimates are poor probably because diffusion reached a high value very early on, after which time the models have little explanatory power over the small changes in S(i,t)/T(i,t) some of which were negative.

6 Conclusion

The objective of this paper was to use the example of steam and motor ship diffusion to test the hypothesis that world diffusion affects domestic diffusion. We presented a simple extension to the classic epidemic diffusion models of Bass (1969) and Mansfield (1961). Our measure of world diffusion is the simple average of steam and motor ship diffusion in a group of 13 countries. We find some evidence for spillovers but the sign of the effect varies across countries. One reason for this may be that because steamships became competitive only gradually as technology improved, information was important only when it concerned the particular market in which the ship owner was operating. In this case, the positive effect that a high level first observation domestic diffusion is expected to have may be countered by the effect of information from abroad from a market more relevant to the particular ship owner. Although we cannot test this effect it suggests that the different signs of the foreign effect coefficient may be due to a failure to adequately separate between the domestic and world diffusion effects.

Another possible explanation for the varying signs of the parameters may be the nonstationarity of the model variables. We argued that nonstationarity in the classic epidemic models is a major unrecognised problem and that any extension that nests within it the Bass or Mansfield models will also suffer from nonstationarity. The question is how to develop a model that does not have a problem with nonstationarity. The logistic transformation that we applied to the diffusion measure passed the stationarity tests more frequently, but not consistently enough. We conclude that the construction of an epidemic model that is satisfactory from a time-series point of view would take a more systematic effort. We will leave this as a topic of further study.

Appendix

A time-series process is said to be strictly stationary if its properties are unaffected by a change of time origin (Verbeek 2000:228). We are concerned with weak stationarity, which implies that the means, variances and covariances of the series are independent of time. If we plot a stationary time series over time we find that it periodically returns to its long-run mean, whereas a non-stationary series typically "wanders off". An example of a non-stationary series is a random walk.

The reason why the stationarity of economic variables is a concern is the spurious regression problem. Let Y(t) and X(t) be two variables generated by independent pure random walks:

$$Y(t) = Y(t-1) + u(1,t) \quad u(1,t) \sim IID(0,\sigma_1^2)$$

$$X(t) = X(t-1) + u(2,t) \quad u(2,t) \sim IID(0,\sigma_2^2)$$

where u(1,t) and u(2,t) are mutually independent. Granger and Newbold (1974) showed that in the OLS regression

$$Y(t) = \delta + \mu X(t) + v(t)$$

the t-ratio on μ is likely to be significant, despite the lack of a causal relationship between Y(t) and X(t). To rule out the possibility of spurious regression we test for the degree of stationarity in our variables, before proceeding with least-squares estimation. If explanatory variables are non-stationary, inferences made using conventional asymptotic theory for least-squares estimation are unreliable.

We now describe Perron's structural break test. First, we allow for a break in both the intercept and trend at time T_B. The alternative is trend stationarity, with a structural break. The appropriate regression is

$$\begin{aligned}\Delta Y(i,t) = \delta + \varphi t + (\mu-1)Y(i,t-1) + \delta_1 DL(T_B) + \delta_2 DP(T_B) + \delta_3 t * DL(T_B)\\ + \sum_{j=1}^{n} c(j)\Delta Y(i,t-j) + u(i,t)\end{aligned} \quad . \qquad (13)$$

Here, Y(i,t) is the variable that is being tested for stationarity. The null hypothesis of non-stationarity is tested using the t-ratio on (μ-1) for which critical values are provided by Perron (1989, Table VIB). DL and DP are dummy variables that are used to captures the changes in the intercept and time trend at T_B. The lag length n is chosen so that the residual u(i,t) is white noise. Beginning with 15 lags, we used the following criteria. If the t-statistic on the longest lag was significant at 5 per cent, that lag length was taken. If the longest lag was significant at

10 per cent, we used the Akaike information criterion (AIC) and the equation standard error to determine whether the lag could be dropped. Lag lengths are reported in the tables below.

When the time dummies were not significant, alternative versions of Perron's test were conducted. If the structural break was not significant, the ADF test with a constant and a time trend was run, i.e. This corresponds to (13) with $\delta_1 = \delta_2 = \delta_3 = 0$. The significance of the time trend φt was tested using critical values from Dickey and Fuller (1981, Tables V and VI) and if insignificant, the time trend was dropped in order to increase the power of the test. To test the null hypothesis in the ADF test, we used critical values from MacKinnon (1991).

We present here the results of the unit root tests referred to in section 5.1. Where the null of a unit root was rejected in the ADF test, results are only reported for this. When no date is indicated for T_B the results also refer to the ADF test. We report the results for same period as the regression results in Section 5.2 unless a unit root was only rejected for the other period, in which case results are reported for both periods.

Table A1. Unit root tests for the Bass and Mansfield model variables

Panel A: $\Delta S(t)/T(t-1)$

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	A		0	0.17	Y	-5.57
Austria	1837-1912	B	1894	7	-0.48	Y	-6.02
Belgium	1837-1931	A		4	0.42	5%	-2.98
Canada	1892-1939	A		0	0.41	Y	-4.28
Denmark	1844-1913	A		0	0.36	Y	-5.55
Finland	1873-1913	A		0	-0.06	Y	-6.16
France	1839-1939	A		2	0.36	Y	-6.52
France	1838-1913	A		9	0.33	N	-2.53
Germany	1850-1938	B	1913	11	-0.03	Y	-3.83
Germany	1850-1913	A		11	0.82	N	-1.41
Italy	1862-1925	A		7	-3.44	Y	-4.58
Italy	1862-1913	A		12	1.58	N	2.06
Netherlands	1846-1913	A		5	-0.41	Y	-4.71
New Zealand	1870-1939	A		0	-0.20	Y	-9.90
Norway	1866-1913	A		2	-0.05	Y	-6.15
Sweden	1865-1913	A		1	0.17	Y	-4.40
United Kingdom	1837-1913	A		0	-0.23	Y	-10.76
United States	1837-1939	A		1	0.63	Y	-5.08
United States	1837-1913	A		2	0.54	5%	-3.15
World	1837-1938	A		2	0.42	Y	-5.81
World	1837-1913	A		0	-0.05	Y	-8.91

Notes to Table A1: "Model" indicates the specific unit root test that was used: A allows no structural break; B allows a single break in the intercept; C a break in the slope; D a break in both intercept and slope. T_B is the date of the structural break. μ-ADF is the estimate of the autoregressive coefficient. Column "reject?" indicates whether a unit root is rejected at 1 per cent (Y), at 5 per cent (5%) or not (N). n is the number of lags. t-ADF is the t-ratio on (μ-1).

Panel B: S(t-1) / T(t-1)

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	B	1919	0	0.82	N	-1.93
Austria	1837-1912	A		5	0.98	N	-1.78
Belgium	1837-1931	A		13	0.95	N	-1.97
Canada	1892-1939	D	1921	0	0.59	N	-3.33
Denmark	1844-1913	D	1871	1	0.86	N	-2.89
Finland	1873-1939	D	1920	0	0.34	Y	-6.94
Finland	1873-1913	A		0	0.64	N	-2.66
France	1838-1913	A		8	0.92	N	-2.79
Germany	1850-1913	A		4	0.96	N	-2.32
Italy	1862-1913	A		10	0.91	N	-1.54
Netherlands	1846-1913	A		13	0.88	5%	-4.00
New Zealand	1870-1939	C	1913	8	0.40	Y	-5.01
Norway	1866-1913	A		0	0.99	N	-0.48
Sweden	1865-1913	A		13	0.81	N	-2.97
United Kingdom	1837-1913	A		1	0.94	N	-2.60
United States	1837-1913	A		11	0.93	N	-1.25
World	1837-1938	A		6	0.97	N	-1.70
World	1837-1913	D	1865	0	0.73	N	-3.57

Panel C: [S(t-1) / T(t-1)]^2

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	B	1919	0	0.55	N	-3.62
Austria	1837-1912	D	1878	9	0.88	N	-3.81
Belgium	1837-1931	A		5	0.96	N	-2.21
Canada	1892-1939	D	1921	0	0.75	N	-2.47
Denmark	1844-1913	D	1871	6	0.96	N	-1.59
Finland	1873-1939	D	1920	5	-0.21	Y	-7.09
Finland	1873-1913	A		2	0.72	N	-2.40
France	1838-1913	A		14	0.89	N	-2.12
Germany	1850-1913	A		6	0.96	5%	-3.87
Italy	1862-1913	A		19	2.39	N	2.17
Netherlands	1846-1913	A		13	0.81	Y	-4.57
New Zealand	1870-1939	B	1913	3	0.76	N	-3.24
Norway	1866-1913	A		15	2.31	N	2.58
Sweden	1865-1913	A		13	0.85	N	-2.32
United Kingdom	1837-1913	A		0	0.99	N	-0.83
United States	1837-1913	A		1	1.01	N	0.81
World	1837-1938	A		12	0.97	N	-3.07
World	1837-1913	D	1865	3	0.95	N	-2.35

Notes to Table A1: "Model" indicates the specific unit root test that was used: A allows no structural break; B allows a single break in the intercept; C a break in the slope; D a break in both intercept and slope. T_B is the date of the structural break. μ-ADF is the estimate of the autoregressive coefficient. Column "reject?" indicates whether a unit root is rejected at 1 per cent (Y), at 5 per cent (5%) or not (N). n is the number of lags. t-ADF is the t-ratio on (μ-1).

Table A2. Unit root tests for the additional variables in the general model

Panel A: [W(t-1) - S(t-1)] / [TW(t-1) - T(t-1)]

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	A		3	0.95	N	-0.64
Austria	1837-1912	D	1865	1	0.70	N	-3.77
Belgium	1837-1931	A		6	0.96	N	-2.09
Denmark	1844-1913	A		0	0.97	N	-2.00
Finland	1873-1913	A		0	0.97	N	-2.13
France	1838-1913	A		0	0.97	N	-2.18
Germany	1850-1913	A		0	0.91	N	-3.00
Netherlands	1846-1913	A		0	0.97	N	-2.09
New Zealand	1870-1939	D	1913	1	0.79	N	-2.93
Norway	1866-1913	A		0	0.82	N	-1.96
Sweden	1865-1913	A		0	0.77	N	-2.61
United Kingdom	1837-1913	A		11	0.98	N	-0.78
United States	1837-1913	D	1864	1	0.74	N	-3.52

Notes to Table A2: "Model" indicates the specific unit root test that was used: A allows no structural break; D allows a break in both intercept and slope. T_B is the date of the structural break. λ is the ratio of the pre-break sample to total sample size. μ-ADF is the estimate of the autoregressive coefficient. Column "reject?" indicates whether a unit root is rejected at 1 per cent (Y), at 5 per cent (5%) or not (N). n is the number of lags. t-ADF is the t-ratio on (μ-1). Canada and Italy are missing from Panel A because since they are not included in W(t) and TW(t) the ADF test on the variable [W(t)-S(i,t)]/[TW(t)-T(i,t)] is not relevant for them.

Panel B: [S(t-1) / T(t-1)] * [W(t-1) - S(t-1)] / [TW(t-1) - T(t-1)]

Country	Period	Model	T_B	n	μ-ADF	reject?	t-ADF
Australia	1876-1925	D	1919	0	0.69	N	-3.39
Austria	1837-1912	D	1878	12	0.85	N	-3.61
Belgium	1837-1931	A		13	0.92	N	-3.26
Canada	1892-1939	D	1921	0	0.69	N	-2.73
Denmark	1845-1939	B	1897	5	0.94	Y	-3.90
Denmark	1844-1913	D	1871	6	0.96	N	-1.84
Finland	1873-1939	D	1920	0	0.25	Y	-8.18
Finland	1873-1913	A		0	0.78	N	-2.79
France	1838-1913	A		6	0.97	N	-1.22
Germany	1850-1913	A		6	0.97	N	-2.75
Italy	1862-1913	A		16	1.42	N	2.61
Netherlands	1846-1913	A		13	0.87	Y	-4.96
New Zealand	1870-1939	A		1	0.91	N	-2.00
Norway	1866-1913	A		11	0.95	N	-1.03
Sweden	1865-1913	A		9	0.96	N	-1.56
United Kingdom	1837-1913	A		0	1.01	N	1.37
United States	1837-1913	D		0	0.98	N	-1.28

References

Alderton, T. and Winchester, N. 'Globalisation and De-Regulation in the Maritime Industry', *Marine Policy* 26, no. 1 (2002), pp. 35-43.

Barton, J. R. ' Flags of Convenience : Geoeconomics and Regulatory Minimisation', *Tijdschrift voor Economische en Sociale Geografie* 90, no. 2 (1999), pp. 142-155.

Bass, F. M. 'A New Product Growth for Model Consumer Durables', *Management Science* 15, no. 5 (1969), pp. 215-227.

Breshanan, T. F. and Trajtenberg, M. 'General Purpose Technologies: "Engines of Growth"?' *Journal of Econometrics* 65 (1995), pp. 83-108.

Broeze, F. J. A. 'The Cost of Distance: Shipping and the Early Australian Economy, 1788-1850', *The Economic History Review* 28, no. 4 (1975), pp. 582-597.

Cohn, R. L. 'The Transition from Sail to Steam in Immigration to the United States', *Journal of Economic History* 65, no. 2 (2005), pp. 469-495.

Comin, D. and Hobijn, B. 'Cross-Country Technology Adoption: Making the Theories Face the Facts', *Journal of Monetary Economics* 51 (2004), pp. 39-83.

Comin, D., Hobijn, B. and Rovito, E. 'Five Facts You Need to Know About Technology Diffusion', in, *Working paper*, (Cambridge, MA., 2006).

Crafts, N. 'Steam as a General Purpose Technology: A Growth Accounting Perspective', *The Economic Journal* 114 (2004), pp. 338-351.

Dickey, D. A. and Fuller, W. A. 'Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root', *Econometrica* 49, no. 4 (1981), pp. 1057-1072.

Geroski, P. 'Models of Technology Diffusion', *Research Policy* 29, no. 4 (2000), pp. 603-625.

Graham, G. S. 'The Ascendancy of the Sailing Ship 1850-85', *The Economic History Review* 9, no. 1 (1956), pp. 74-88.

Granger, C. W. J. and Newbold, P. 'Spurious Regressions in Econometrics', *Journal of Econometrics* 2, no. 2 (1974), pp. 111-120.

Greenhill, B., *The Merchant Schooners* (Newton Abbot, 1968).

Gruber, H. and Verboven, F. 'The Diffusion of Mobile Telecommunications Services in the European Union', *European Economic Review* 45 (2001), pp. 577-588.

Harley, C. K. 'The Shift from Sailing Ships to Steamships, 1850-1890: A Study in Technological Change and Its Diffusion', in D. N. McCloskey ed., *Essays on a Mature Economy: Britain after 1840*, (London, 1971).

———. 'Ocean Freight Rates and Productivity, 1740-1913: The Primacy of Mechanical Invention Reaffirmed', *The Journal of Economic History* 48, no. 4 (1988), pp. 851-876.

Howells, J. 'The Response of Old Technology Incumbents to Technological Competition - Does the Sailing Ship Effect Exist?' *Journal of Management Studies* 39, no. 7 (2002), pp. 887-906.

Knauerhase, R. 'The Compound Steam Engine and Productivity Changes in the German Merchant Marine Fleet, 1871-1887', *The Journal of Economic History* 28, no. 3 (1968), pp. 390-403.

Kumar, V. and Krishnan, T. V. 'Multinational Diffusion Models: An Alternative Framework', *Marketing Science* 21, no. 3 (2002), pp. 318-330.

MacKinnon, J. G. 'Critical Values for Cointegration Tests', in R. F. Engle and C. W. J. Granger eds., *Long-Run Economic Relationships: Readings in Cointegration* (Oxford, 1991), pp. 267-276.

Mansfield, E. 'Technical Change and the Rate of Imitation', *Econometrica* 29, no. 4 (1961), pp. 741-766.

Mitchell, B. R., *International Historical Statistics* (Basingstoke, 4 edn, 1998).

———, *International Historical Statistics* (Basingstoke, 5 edn, 2003).

Perron, P. 'The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis', *Econometrica* 57, no. 6 (1989), pp. 1361-1401.

Samstag, T. and Joshi, R., *Norwegian Shipping - the Past, Present and the Future* (Oslo, 2005).

Strang, D. and Soule, S. A. 'Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills', *Annual Review of Sociology* 24 (1998), pp. 265-290.

Verbeek, M., *A Guide to Modern Econometrics* (Chichester, 2000).