DISCUSSION PAPER SERIES

No. 5890

THE DYE FAMINE AND ITS AFTERMATH: KNOWLEDGE DIFFUSION AND ENTRY

David Genesove

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

ISSN 0265-8003

THE DYE FAMINE AND ITS AFTERMATH: KNOWLEDGE DIFFUSION AND ENTRY

David Genesove, Hebrew University of Jerusalem and CEPR

Discussion Paper No. 5890
October 2006

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: David Genesove

CEPR Discussion Paper No. 5890

October 2006

ABSTRACT

The Dye Famine and its Aftermath: Knowledge Diffusion and Entry*

A firm that introduces a new good enjoys monopoly profits for some initial period of time. What happens subsequently depends upon the relative strength of knowledge diffusion and increasing dominance. The first effect enhances challengers' ability to develop the product, erodes the incumbent's monopoly power, while the second, which concerns the relative net cost of the incumbent to challengers in production, strengthens it. This paper exploits the near total disruption of imports of German dyes to the United States during World War I and the immediate post-War period, and the subsequent re-entry of the Germans to the market, to separately estimate the first effect. The results show that while (a) the probability of a dye was imported in 1913-1914 bore no relation to its year of discovery, (b) the probability it was produced in 1917 by the new American manufacturers was greater by one and a half percent per year, the earlier the year of discovery. Coupled with the estimated semi-elasticity of the probability of production with respect to the volume of imports in 1914, and assuming prospective profits were proportional to that volume, one obtains that every additional year since discovery decreased the expected cost of developing the dye by 19 to 25%. The paper shows, additionally, that after German imports were able to resume, the probability of a dye being imported in 1923, given that the Americans were already producing it, was independent of the year of discovery – implying that the discovery year is an appropriate proxy for the amount of development relevant knowledge that had diffused through the industry.

JEL Classification: L65, O31 and O33
Keywords: dyes, increasing dominance, innovation and knowledge diffusion

David Genesove
Department of Economics
Faculty of Social Sciences
Hebrew University of Jerusalem
Mount Scopus
Jerusalem 91905
ISRAEL
Email: genesove@mscc.huji.ac.il

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=121880

* I am grateful to Tony Travis for tutoring me in the history and chemistry of the dye industry, Kathryn Steen for pointing me to a number of important documents, Ashish Arora for his comments, Pierre Regibieau, Joachim Henkel and Julie Holland Mortimer for acting as discussants at CEPR and NBER conferences, and participants there and at the European University Institute, Bar-Ilan, the London Business School, Hebrew University and the LSE. Support from the Science Technology and the Economy Program (STE) at the Samuel Neaman Institute for Advanced Studies in Science and Technology is gratefully acknowledged.

Submitted 15 September 2006

Introduction

Incumbency generally derives from an informational advantage over other firms. At least initially, the incumbent knew how to make the product better than others – who may not have known how to make it at all. Arrayed against this initial advantage is knowledge diffusion. Knowledge leaks out of the firm, or develops independently elsewhere in the commercial realm, thus threatening the initial dominance.

Estimating the speed of knowledge diffusion, say through the entry of new firms or the erosion of early firms' market shares, is confounded by a second mechanism. A number of theoretical papers have argued that many types of competition exhibit increasing dominance, in which firm the firm's dominant position is perpetuated through the complementarity of investment and market presence. That complementarity provides an advantage to the incumbent or leading firm over rivals, potential or current, and so leads to a decreasing likelihood of entry, or an increasing market share, over time.[1] Thus the persistence of dominance, or its deterioration, is the outcome of a race between knowledge diffusion and increasing dominance. What is observed is the net effect of the two forces, and not each separately.

It would thus be useful to observe an industry in which the increasing dominance effect is absent – say, by removing the incumbent firms from the market. In that case, entry would be determined solely by the attributes of the entrant and not those of the incumbent firm, and so the knowledge diffusion effect would be isolated.

The opportunity to observe how potential entrants behave when the incumbent is removed from the market is surely a rare one, but the synthetic dye industry in the United States during World War I and the years that followed it provides precisely that opportunity. This industry was heavily dominated by German companies before the war, which was to cut off many countries from this supply. In particular, the United States, although initially a non-combatant, was denied German dyes due to the British naval blockade. The resulting dye famine, as it was called, induced domestic firms to enter the industry. This situation obviously continued after the U.S. entry into the war in the summer of 1917. It ended with the re-entry of German firms a couple of years

---

[1] The simplest mechanism that generates complementarity is the greater cost saving that a firm with a greater output has for a variable cost reducing investment, that itself incurs a fixed cost, as in Klepper (1996), which arises whatever the market structure. For models in which the complementarity arises for additional strategic reasons, see, e.g., Gilbert and Newbery (1982), Cabral and Riordan (1994), and a general treatment under myopic behaviour, Athey and Schmutzler (2001).

after the end of hostilities, by which time some of the U.S. firms had established a sufficient foothold in the industry, made surer by extremely high tariffs.

This paper asks which previously imported dyes the American firms succeeded in producing when German imports were cut off. In particular, were the American firms relatively more likely to produce the newer or older dyes? The question is relevant under the maintained assumption that the older the dye (i.e. the earlier the year of its discovery) the longer the time that knowledge about its composition and its manufacture has to diffuse through the economy. Furthermore, by comparing the effect on production incidence of a dye's age to that of the log 1914 import quantity, proxying for prospective profits, we can infer the rate at which the entry cost necessary to compensate for the lack of complete knowledge declines with the dye's age.

The use of the import quantity to proxy for profits misses any variation in the price-cut margin. The resulting biases are assessed and attempts are made to reduce them through use of demand and cost proxies. Of course, newer dyes may be more or less complicated to develop than older ones, so there will be a need to control for that as well.

Since the war eventually ended and the Germans returned to export to the United States, there is an opportunity to observe the behaviour of entrant and incumbent after the entrant has sunk its development cost. By asking which dyes the Americans continued to produce after the re-entry of the German firms, we can determine whether knowledge diffusion and increasing dominance, on net this time, is relevant to post-entry competition as well. The fact that not all dyes were developed by the American firms in the years between the British blockade and the Germans' return to the market introduces a selection bias, which is dealt with by restricting the sample to those dyes produced by either the Americans or imported from abroad, but not both. The approach is analogous to the use of conditional logit in panel studies.

The dye industry has a number of attributes that make it an attractive industry to study, aside from the Dye Famine itself. The chemical industry as a whole is a striking example of the ability of firms to recover their dominant position despite massive negative shocks to their physical capital and intellectual property. Cantwell (1995, 2000) has shown how Germany remained dominant in U.S. patenting in this industry despite the confiscation of its patents in both world wars, the copying of its firms' internal documents and the destruction of many of their factories during the second, which points to the crucial roles played by human and organizational capital in firms' success. Arora (1995a, 1995b) has used this industry to show the limited value to firms of intellectual property, without the accompanying human capital.

There are more prosaic reasons for studying this industry. From early on, dyes, as chemical compounds, were thoroughly categorized. The first to do so was Gustav Schultz, and it is the fifth edition (Schultz, 1914) of his work that is used here. The Schultz category corresponds to a unique chemical compound. Because of possible differences in concentration, of which more later, it is not necessarily a precisely defined product; however it is a set of very nearly perfect substitutes.

An additional advantage to studying this industry is that data on U.S. production and imports were comprehensively gathered by the U.S. government. In aid of the deliberations over the 1916 tariff legislation, the Commerce Department collected and published figures on the pre-war quantity of dye imports (and, for a subset of them, their value), by the Schultz category (Norton, 1916).[2] Then, as required by that legislation, which provided for the eventual removal of the tariff should U.S. dye production exceed 60 percent of consumption, the Tariff Commission collected and reported production and imports, also by the Schultz number, yearly from 1917 on. Due to censoring according to the usual three firm rule, quantity figures are available only for dyes produced by at least three firms, which helps explain why this work is concerned only with the incidence of production, and not its extent.

Also, because the dye industry is generally regarded as the first high tech industry, with scientists controlling the firms by 1914, and often much earlier, (e.g., Liebenau, 1992, Marsch, 1994), dedicated R&D laboratories, a high level of academic-industry collaboration and a Nobel prize awarded in 1905 to one of its consultants (Adolph Baeyer), and, because the industry is closely connected to two other important industries - the explosives and pharmaceutical industries - , there is a vast business history literature about it, which is an advantage to an econometric study of any industry.

Section II: Pre-History

---

[2] Steen (1995a, p. 118) notes that Norton "assembled the data from confidential [customs] invoices completed on the dyes' arrival to the country. Norton's method generally gave the statistics veracity, although manufacturers mistrusted his figures for the dyes such as indigo, which the tariff laws excluded. The invoices for excluded chemicals proved to be much less reliable." She fails to note how those figures were "proved" to be so. I supplement the data from Norton (1916) with additional price figures from Norton (November 1916) for about 30 dyes, and for calculation of the HHI, Norton (1922), which provides firm level quantities.

The synthetic dye industry is usually dated to 1856, the year of Perkin's discovery of the first aniline dye, which yielded the colour mauve. Synthetic dyes are produced from intermediates, themselves the by-products of tar-crudes from coal ovens. Perkin's discovery and profitable commercialization of this dye prompted other chemists in the UK and elsewhere to research this area, which led to the discovery of other dyes and manufacturing processes.

Although the industry was initially dominated by the UK and France, within fifteen years it was clearly dominated by Germany. This shift was embodied in the homeward migration of a number of German chemists who had been residing in England, among them was H. Caro, who simultaneously with Perkin in 1869 developed the first of the alizarin dyes, the second major class of dyes to be developed, as well as A. Hoffman, Perkin's initially sceptical dissertation advisor. Those firms that were at all successful elsewhere, particularly in the United Kingdom and the United States, were generally founded by German emigrants.

Most economic historians (e.g., Beer, 1959, and most recently Murmann, 2003) have attributed Germany's comparative advantage in this industry to its universities and their close working relationship with the industry. The relatively greater esteem given to industrial research in Germany is also cited. Some role seems also to have been played by the patent systems in the various countries, with the French system providing a very broad patent that covered the dye whatever the process by which it was produced, and the British system providing little prior examination, and so leading to massive patent litigation (e.g., Hohenberg, p. 71). The German states, in contrast, had essentially no patent law[3] allowing the new firms there to grow by copying the British and French innovations (Beer). By the time patent legislation was introduced in the newly formed Germany in 1877, the surviving German firms were themselves developing new dyes and the more stringent conditions for patents helped protect them and stimulate industry research (Marsch, 1994). The industry was sufficiently developed in Germany by this point that it did not shift elsewhere, not even across the border to Switzerland which boasted a smaller dye industry itself, but lacked patent law.

Of most interest to us is the failure of the industry to develop in the United States. Taussig (1922) blamed the absence of economies of scale, so central to

---

[3] Although certain German states did have a patent law, the patent protection did not cover imports other German states (Murmann, 2003).

American success in other industries, and the abundance of skilled, factory-level labour in Germany, but the dearth of trained organic chemists in the United States was surely crucial as well. Arora and Rosenberg (1998), also note the use of the beehive ovens in the United States, which did not provide the coal-tar by-product necessary for production of the intermediates.

Figure 1 shows the distribution of dyes by the discovery year[4] for the population of dyes discovered in 1856 or later. The line marks 1875 which serves as the cut-off point for the sample used in the rest of the paper, given the relatively small number of discoveries before that date. The distribution is relatively spread out, with a single peak in the early 1890s. Part, but not all, of the sharp fall-off after 1910 surely reflects Schultz' ignorance in 1913/4 of the very most recent discoveries; *The Color Index*, (Rowe, 1924), shows a more moderate fall-off, with the number of discoveries post-1910 at only a third or fourth of the immediately preceding years.

German firms were extremely dominant in the industry. In 1913, on the verge of the war, they were producing between 83 to 88 percent of the dyes consumed in the world. The only other net exporter was Switzerland, which produced about six percent of world production. In contrast, American firms produced only about two percent of world dyes, or about 13% of U.S. dye consumption (Reader, 1970, p. 252 and 258). There were seven American firms producing dyes before World War I. By far, the largest was Schoellkopf, which produced 106 different dyes. No other firm, including Bayer, the only German subsidiary operating in the United States, produced more than fifteen. Altogether, U.S. firms produced some 130 different dyes, compared to the 922 listed by Schultz and the over 500 imported from abroad. The number 130 vastly overstates U.S. presence in the industry, however, for the U.S. firms were essentially "assemblers"[5] and not "producers". According to Haynes (1945, Vo. 3, p. 213), of the seven dye U.S. dye producers, only Schoellkop made any of its own intermediates, and, in addition, "[o]nly one American chemical company, Benzol Products, made any

---

[4] For a few dyes, Schultz lists more than one year of discovery, presumably in cases where there were either independent discoveries or in which there were different processes. In such cases, I use the earlier year. Schultz dates some of the compounds earlier than the traditional starting point of 1856.

[5] Joseph H. Choate, Jr., counsel for the Chemical Foundation and American Dyes Institute, used the same terminology when he testified that "Our industry was a mere assembling industry operating … on German intermediates. We imported things almost finished from Germany and turned them into finished dyes here by final processes which are often very simple." (*Dyestuff Hearings*, 1919, p. 90.) Likewise Norton (November 1916) writes, "The American manufacture was confined almost entirely to the "assembling" into finished dyes of coal-tar intermediates imported from Europe, chiefly from Germany." Little (March 1915) and Hesse (November 1915) also use this term.

pretense of being a real producer of coal-tar intermediates and its claims were not very impressive.”. When German imports were cut off, Schoellkopf was forced to cut back its dye production from the 106 to fifteen categories only, most of which were black or blue (Steen, 1995, p. 124). Germany’s dominance in the industry is also plainly seen in its firms’ share of U.S. dye patents. Eighty one percent of the 1444 U.S. dye patents issued between 1900 and 1917 and listed in Doyle (1926) were assigned to German firms. Another ten and a half were assigned to Swiss firms. A mere four and a half percent went to American firms.

An examination of the determinants of pre-war imports provides a useful baseline for analysing the incidence of American production after the Dye Famine. Figure 2a shows the fraction of dyes imported in fiscal year 1914 (July 1, 1913- June 30, 1914) by the year of discovery, taken from Norton (1914). The regression line, whose estimates are shown in the first column of Table 2, overlays the figure. It is negatively sloped, but by only a statistically insignificant 1/10 of a percent decline by year (indicating a 3.8% difference over the range of the years in the sample). Thus there was no differential tendency for earlier or later dyes to be imported.

The remaining columns of Table 2 consider the effect of adding additional demand and cost proxies. Column (2) adds the count of countries among Germany, the U.S., England and France, in which at least one patent had been taken out in the dye category, according to Schultz (1914), as an indicator of profitability.[6] It is significant, but does not affect the relationship between the year of discovery and import incidence. Column (3) adds dummy variables for the 16 dye classes, listed in Table 1. The dye classes indicate a shared common tar-crude or even basic intermediate, and so constitute the fundamental technological grouping of the dyes. The F-test statistic is 2.43, with a p-value of .002. Column (4) adds the Schultz Number itself (normalized to vary between 0 and 1), which is insignificant. Column (5) adds 15 (non-exclusive)[7] dummy variables for the dye’s colours. Column (6) adds (non-exclusive) dummy variables for the material (cotton, wool and silk) to which the dye may be applied.[8] Both sets of variables are highly significant. Finally, column (7) includes all the aforementioned variables. Each (set of) variable(s) remains significant (or

---

[6] For the use of the count of countries with patents, see, for example, Putnam, (1997).

[7] Dyes can often be altered slightly in their application to produce different colours.

[8] The source of the data for both sets of variables is *Colour Index,* 1924 Schultz, 1914, also presents this information, but in German.

insignificant, in the case of the Schultz number). Throughout, the coefficient on the discovery year remains essentially unchanged.

These results are important in establishing that the discovery year is not proxying for demand or cost in our later analysis of U.S. production in 1917. The incidence of importation should be a function of U.S. demand (relative to the past, expected world demand, that helped determine the set of dyes available) and cost. We see that in the high significance levels of the exclusion tests for indicators of colour, application material, patents and dye classes. In contrast, the discovery year is clearly not a determinant of importation, indicating that it proxies for none of the major demand or cost factors that we observe.

Section III: The Dye Famine, the Entry of American Firms and Technology Transfer

Soon after the start of hostilities, Britain's navy blockaded Germany.[9] This did not immediately cut off the supply of all dyes to the United States, as there were domestic and, more importantly, foreign stocks to draw from. The domestic stocks apparently sufficed for a few months only,[10] but the availability of the foreign stocks and the general impression that the war would not last very long meant that investments in dye production were not seen as profitable at this stage.

But the foreign stocks eventually ran out. Indigo, seven to eight million pounds of which had been imported annually between 1909 and 1914, with a German share never less than 92 percent, continued to be imported in 1915 at 8 million pounds, albeit with a German share of only 86 percent, the difference being made up primarily by European and South American countries. But in 1916, imports fell to six and a half million pounds, none of which came from Germany. Half of the imports now came from (German produced stocks in) China; the U.K. (where a German owned indigo plant had been confiscated) and India made up the rest of the German deficit. In 1917, the foreign stocks were depleted, and indigo imports had fell to half a million. For other dyes, the shortage came earlier. Annual imports of Alizarin dyes had varied between three to eight million pounds between 1909 and 1914, with six million pounds

---

[9] There was an immediate self-imposed German embargo, but it was quickly removed. Twice German dyes were delivered by submarines that evaded the British blockade. But the submarines carried a relatively small supply and were essentially a publicity stunt (Steen).

[10] According to one importer, there were only 60 days worth of domestic stocks. Little (March 1915) suggests that textile mills had five months' stock.

(all but six percent from Germany) imported in 1915. In 1916, German imports went to zero, and their, being no substitute stocks, imports during the remaining war years, were less than one percent of the 1915 level.[11]

The so-called Dye Famine, was characterized not only by low imports and consumption but, naturally enough, high prices, and substitution to natural dyes. Figure 3, taken from Jones and Cassebeer, 1919, shows the evolution of an index of dye, intermediate and coal-tar crude prices over this period. Relatively stable before the war, it doubles at the start of the war and reaches a peak of more than 7 times its pre-war real value at the end of 1915, before returning to its real pre-war value at the end of 1919.[12] (By April 1915, the stocks were drawn down) Figure 4, from the same publication, shows that the price of natural dyes, the only substitute, rose dramatically as well.

The profit opportunities were obvious, and many U.S. firms entered the dye industry. By December 1915, fourteen firms were operating, by February 1916, there were sixteen firms, and by May 1916, there were twenty four;[13] by the time of the 1917 Census, some 81 firms overall were manufacturing synthetic dyes. The largest investor of them all was Du Pont, which had already been operating in the related explosives industry. But notwithstanding its eventual 21 million dollar investment, Du Pont was to run second in market share to National Aniline, the result of a merger of a number of vertically related firms already operating in the generally industry, including Schoellkopf.

The major impediment for the U.S. firms was their lack of know-how in making dyes. This required the firms to undertake a variety of technology transfer strategies. The first, obvious one was to turn to the chemical literature, comprising textbooks, academic and trade journals, and patents. Unfortunately for the firms, each of these sources proved inadequate. According to Haynes, cited by Travis (2004, p, 41), the textbooks were “ten years behind current chemical plant practises”, and the journals never described the necessary processes exactly. This does not mean that the literature was valueless. Hounshell and Smith (1988) relate, for example, how Du Pont relied on a footnote from the *Journal fur pratische chemie* to substantially improve the

---

[11] United States Tariff Commission, 1918. Nonetheless, Norton (February 1916) states that no dyes were imported from Germany past March 15 1915, and only small amounts were imported from Switzerland.

[12] The figures presented in United States Tariff Commission, 1918 (p. 8) are broadly consistent, showing an increase of 3.5 times of the (weighted) average price in 1916 over 1913.

[13] See Norton (February 1916), Norton (March 1916), and Norton (May 1916).

design for its diphenylamine plant, an important explosive intermediate. But the chemical literature by itself did not give the firms the necessary capability to produce the dyes.

Particularly inadequate were the patent descriptions. These were incomplete as well, with vital information on catalysts, as well as optimal temperatures, pressures and timing missing.[14] According to the admittedly pro-American bias of Haynes (Vol. 3, p. 214), the patents "had deliberate gaps and were deceitfully misleading". Further complicating matters was that it was not always clear which patents were relevant to which dyes. There were even claims that some patents, so called 'evasion patents', had been taken out only to mislead competitors (Hounshell and Smith, p. 89), and one Du Pont executive went so far as to assert that it took almost as long to determine the match between a patent and a dye as to discover the dye in the first place – surely an exaggerated claim, yet still an indication of the difficulties that the American firms faced in using the patent literature.[15]

Technology was acquired by way of human capital as well. During the war years, this was limited on the research end to the small pool of American chemists who had trained or worked in Germany. Calco, for example, used a Yale chemistry professor with such a background (Travis, 2004). But human capital was important elsewhere in the firm as well. The American firms hired the local marketing agents for the German firms, who would have had important knowledge about demand, and buyers' identities (Hounshell and Smith, p. 82; Steen, 1995a, p. 25, 1995b); those hired by Du Pont brought with them samples of all of the dyes that Badische, a leading German dye manufacturer, had exported to the U.S. Manufacturing skill was also important; Calco's plant manager, for example, had worked in Hoffman-La Roche in Germany (Travis, 2004).

After the war, the way was paved to import German researchers. In 1919 Du Pont tried to hire the leading German dye expert, Rene Bohn (inventor in 1901 of the last new class of dyes, indathrene). Although Bohn found the invitation appalling, and so declined it, some ten more junior chemists took up the offer the following year.

---

[14] These difficulties were demonstrated dramatically in the *USA vs. Chemical Foundation* case, when the judge instructed a young chemist, on the eve of his wedding, to attempt to produce a dye overnight on the basis of the patent description. The chemist failed to produce the dye with sufficient purity (Steen, 2001). How long it would have taken him to get it right, if at all, can not be known.

[15] Japanese firms were apparently also unable to replicate the German dyes based on the patent descriptions (Kudo, 1994) . On the limitation of technical information without accompanying human capital, see Arora (1995a and 1995b).

One received $25,000/year, about $275,000/year in 2004 dollars and a tremendously large salary at the time, and about 50 percent more than the highest earning chemist in Bayer before the war.[16]

A third method was to acquire the necessary knowledge directly from another firm. Du Pont led the way at the end of 1916 by entering into an agreement with the UK firm Levinstein. A pre-war dye firm, founded by a German émigré, it had gained additional knowledge about dye production upon acquiring the Hoechst plant that the UK government had earlier confiscated from its German owners. Du Pont paid Levinstein £25,000 a year for 10 years, in return for all its knowledge on dye production. Several Du Pont employees travelled to England for two months, and returned to write a number of reports, including a 400 page memo containing the 'recipes' for various azo dyes (the largest class of dyes) and intermediates in February 1917. There were at least two more trips in the following two years. Du Pont also built an indigo plant on the plans of the Hoechst plant.[17]

Experimentation was clearly a necessary complement to these various knowledge sources. As Cohen and Levinthal (1989) argued, the ability to absorb others' R&D advances requires an in house capacity of one's own. Hiring a knowledgeable chemist was not enough (Hounshell and Smith), and so the various U.S. firms established large research and development laboratories. Nearly nine percent of the employees of the 190 firms producing either dyes or their intermediates in 1917 were either chemists or engineers. 104 of these firms reported "a separately organized research laboratory for the solution of technical problems in the manufacture of their products and the discovery of new products", and reported spending very close to two and a half million dollars on those labs, or about 3.6% of dye sales.[18] The Tariff Commission authors suspected that more firms conducted research, but "did not keep their books in such a way as to show separately the cost of research" (United States Tariff Commission, 1917). The yearly expenditure rose to slightly over four and a half million dollars in the next year (United States Tariff Commission, 1917).

---

[16] The highest ranked chemist earned 71,727 marks in salary and royalties, Murmann (2003, p. 154) ., citing Citing Meyer-Thurow. I use the 1913 exchange rate of 4.2 marks to the dollar, from Officer, 2005.

[17] Murmann, 2003, offers a couple of 19th century examples in which individuals advertised offers to sell or buy dye recipes.

[18] By comparison, the NSF estimate of the R&D to sales ratio for the much larger category of Industrial Chemicals (1967 SIC codes 281-282), to which the dyes, crudes and intermediates (1967 SIC code 2815) belonged, falls from 5.1 in 1963 to 3.2 in 1974, more or less linearly.

Notwithstanding these various efforts, the Americans were clearly unable to fully replace the dyes that had been cut off from Germany. Table 1 shows the joint distribution of the incidence of importation in 1914 and American production in 1917. Of the 922 dye categories in Schultz' 1914 edition, 835 are ascribed discovery dates that are 1875 or later. Of these, 515 were imported in 1914; only about a quarter of those were produced by the Americans in 1917. Of those dyes not imported in 1914, very few – 20, or some six percent - were subsequently produced in the U.S. This shows stability in demand between the pre-war and war and post-war periods and so provides additional support for the use of pre-war imports as a proxy for demand in the later period.

The inability of the U.S. firms to fully replace the foreign made dyes, even after 1922, when tariffs were raised to 60 percent, led to the last type of technology transfer: ownership. In 1924, CIBA, Geigy and Sandoz, the three major Swiss dye manufacturers, jointly purchased Alt and Wiborg, a major U.S. dye manufacturer. In the most striking case, the leading German firm, Bayer, which had seen its American subsidiary confiscated by the U.S. government and auctioned off, entered into a joint venture in 1924 with Grasselli, which had purchased the Bayer dye patents and plants from the high bidder; Bayer's contribution to the joint venture in the agreement was, for the most part, its technical knowledge. A few years later, IG Farben, the merger of Bayer and all the other major German firms had merged, bought the joint venture outright (Schröter, 1986).

We turn now to consider which dyes the American firms chose to, and were able to, replace.

Section IV: A Model for the U.S. Development Decision

(a) The Model

Consider a dye that was imported into the United States in 1914. Following Bresnahan and Reiss (1990, 1991), we can argue that at least one U.S. firm will supply that dye if and only if monopoly profits (π_i) exceed development costs (F_i).

Let log development costs depend linearly on the number of years since the dye was developed, and so on the year of development, (Y_i), and an independent, median zero error term (f_i),

(1) $$\ln F_i = f_0 + \beta Y_i + f_i$$

Assume further that post-development log monopoly profits equals the 1914 import log-quantity ($q_{1914,i}$) plus an independent, median zero error term (e_i)

(2) $$\pi_i = e_0 + q_{1914,i} + e_i$$

That assumption is motivated below. One thus expects to see a dye produced in the US in 1917 if and only if

(3) $$f_i - e_i \leq [e_0 - f_0] - \beta Y_i + q_{1914,i}$$

and so with probability $H([e_0 - f_0] - \beta Y_i + q_{1914,i})$, where H is taken to be some median-zero distribution, known up to scale. Note that the assumption that post-development monopoly profits are proportional to the import quantity implies that all the parameters are identified. In particular, the estimate of β will be given by minus the ratio of the estimate of the coefficient on the discovery year to the coefficient on the 1914 import quantity, itself an estimate of the inverse of the scale parameter. The last is so, however, only because we have assumed a deterministic process: one pays F to develop a dye with certainty. If there is a constant probability of success, ρ, then the probability of production in 1917 is $\rho H([e_0 - f_0] - \beta Y_i + q_{1914,i})$. β is still identified, as before, but the scale parameter is not if H is the uniform distribution, as will be assumed below.

To motivate the assumption that post-development monopoly profits are proportional to the 1914 import quantity, consider the demand function

(4) $$S_i[sign(\gamma)(\alpha_i - P)]^{\gamma}$$

with either $\alpha_i, \gamma > 0$ or $\alpha_i, \gamma + 1 < 0$. This demand function encompasses a number of commonly used specifications such as the linear ($\gamma = 1$), exponential ($\alpha, \gamma \to \infty, \alpha / \gamma$ finite), log-linear ($\alpha = 0, \gamma < 0$) and inelastic up to a common reservation price ($\gamma \to 0$)

specifications, and is the unique form that generates a monopoly price that is linear in a constant marginal cost (see, e.g., Genesove and Mullin, 1998). It is assumed that S_i varies independently of the other parameters, and that γ does not vary across dyes.

A monopolist facing such a demand curve will set the following price and output

(5) $P_i = [\alpha_i + \gamma c_i]/(1+\gamma)$

(6) $Q_i = S_i Z_i^{\gamma}$

and earn profits

(7) $\Pi_i = \gamma^{-1} sign(\gamma)(\gamma/1+\gamma)^{\gamma+1} S_i Z_i^{\gamma+1}$

where we define $Z_i \equiv sign(\gamma)(\alpha_i - c_i)$, and c_i is a constant marginal cost.

Thus, assuming a foreign monopolist exporting to the U.S. in 1914, the log of the import quantity is an imperfect proxy for the log profits of a U.S. monopolist in 1917:

$$\begin{aligned}\pi_i &= s_i + (\gamma+1)z_i^U + e_i = (q_{1914,i} - \gamma z_i) + (\gamma+1)z_i^U + e_i \\ &= q_{1914,i} + z_i + (\gamma+1)(z_i^U - z_i) + e_i \end{aligned} \tag{8}$$

where small letters indicate the log of the variable, superscript U indicates that the 1917 US marginal cost is substituted for the 1914 foreign marginal cost, and the constant term is ignored. e_i can now be interpreted as the difference between the *horizontal demand shifter* faced by a potential U.S. monopolist in 1917 ($s_i + e_i$) and that faced by the foreign monopolist in 1914 (s_i).

(b) Biases from unobserved price-cost margins

Log 1914 imports differ from prospective 1917 profits by the log price-cost margin z and a term that arises from the difference between foreign and U.S. marginal costs. Were α_i (the *vertical demand shifter*), costs c_i^U and c_i, and thus log-margins z_i and z_i^U, constant across dyes, then 1914 imports would be a perfect proxy for potential profits. It is unreasonable to suppose that, however, and as log-imports is a linear combination of s and z, its estimated coefficient will be biased. That bias will be small if the vast majority of the variation in imports stems from variation in the horizontal demand shifter s. But as we can not be certain that such is the case, we will need to assess the consequent biases, and try to minimize them by employing proxies for the additional terms in (8). In assessing the bias, we will assume that s_i is independent of α_i, c_i^U and c_i

Consider, then, the first additional term: z_i. Noting that $q_{1914,i} = s_i + \gamma z_i$, it is clear that this term imparts a positive (negative) bias to the estimated coefficient on imports when γ is positive (negative).

Next consider the second term, $(\gamma + 1)\{z_i^U - z_i)$. This can be approximated by $(\gamma + 1)(1 - \phi_i)c_i / (\alpha_i - c_i)$, where $\varphi_i \equiv c_i^U / c_i$ can be interpreted as the relative yield. We consider the effect of variation in the two varying factors of this term, $1 - \varphi_i$ and $c_i / (\alpha_i - c_i)$, separately, which is equivalent to assuming a first order Taylor series approximation of the term:

- $1 - \varphi_i$: variations in the relative yield surely stem for the most part from the inadequacies of American production and so are irrelevant for the German monopolist decision in 1914, thus imparting no bias.
- $c_i / (\alpha_i - c_i)$: since either an increase in the numerator or a decrease in the denominator will decrease (increase) z_i if γ is positive (negative), while the term's "coefficient", $(1 - \varphi_i)$, is negative (due to the Germans' marginal cost advantage[19]),

---

[19] All sources make clear that the Germans faced a lower cost of production than the Americans. Schröter (1986, p. 180) writes that specialized dyes cost two to three times as much when produced in a post-ear German subsidiary in the U.S. than in Germany, but that the factor fell for bulk dyestuffs. In detailed testimony before the U.S. Congress in 1908, Schoelkopf claimed that U.S. labor, building and machinery costs were and percent greater than German costs, respectively (Hesse, August 1915). Taussig (1922) attributed the cost differential to the greater availability of skilled, technically educated

we should expect variations in this factor to impart a positive (negative) bias to the estimated coefficient on imports if γ is positive (negative)[20].

Taken together, these terms lead us to expect the estimated coefficient on imports to be positively (negatively) biased if γ is positive (negative), and so to underestimate (overestimate) the magnitude of the knowledge diffusion effect. Thus without knowing the sign of γ we can not sign the bias, and so it is doubly important to somehow control for the unobserved margin.

Including product attributes to proxy for the vertical demand shifter α_i and the production cost terms might serve to eliminate the bias. The demand for a dye is determined primarily by its colour, and the types of clothing to which it may be applied (wool, cotton, silk, etc.), its fastness and the method of application. Production cost is determined by the intermediates used and the nature of the accompanying chemical procedures.

Prices can also be used as a proxy. As equation (5) shows, it is a combination of both vertical demand and cost parameters. Nonetheless, it should prove a useful proxy if only one of α or c varies much. For example, if demand is iso-elastic ($\gamma < 0, \alpha = 0$), price will be a perfect proxy for the price-cost margin z, leaving aside the variation in the degree of concentration, which is discussed below.

Finally, a case can be made that cost should not even be considered a missing variable. American firms would have had limited knowledge about costs when deciding whether to attempt to develop the dye or not. The imported dyes were, of course, made outside of the United States, and their costs of production would have been trade secrets. Fragmented quantity and price information, on the other hand, would have been partially known to the import agencies, whose employees went to work for the new dye manufacturers, and to the dye buyers, that is the textile mills and finishers. With the publication of Norton (1916), all the information on import quantity and prices that is used here, would have been known to the dye firms.[21] The

---

workers in Germany, as well as dye production's lack of substantial economies of scale, the source of U.S. advantage in other industries.

[20] The sign of the bias is the sign of $\gamma(\gamma+1)Cov(z_i, z_i^U - z_i)$ Recall that $(\gamma+1)\gamma$ is always positive.

[21] An early, more detailed draft was shown at the Chemical Exhibition in September 1916 (Steen, 1995a, p. 118). A second draft was published in the *Journal of Industrial and Engineering Chemistry* two months later (Norton, November 1916), with some additional price information, but lacking the smaller dyes, compared to Norton (1916).

crucial question is whether, based on the information they had in hand, the U.S. firms would have projected potential monopoly profits to be proportional to import quantity.

(c) Bias from unobserved (chemical) concentration levels

Differences in the degree of concentration in which dyes were sold will also bias the estimated coefficient on the log of imports. Although two brands of dyes may have had identical production costs and demand parameters, one might have been sold in more concentrated form than the other, and so at a lesser quantity. The 1923 Census report notes that these differences in concentration are to be found even within a dye category. Assuming the (unobserved) degree of concentration to be uncorrelated with the cost and demand parameters (defined in terms of a 100% "pure" product) implies that the quantity we observe is a noisy estimate of the "true" quantity that will cause the estimated coefficient of imports to suffer from the classical errors-in-variable downward bias. Letting h denote the degree of concentration, this bias reduces the magnitude of the estimated coefficient by $Var(h)/Var(q_{1914})$ in the bivariate regression case.

Fortunately, we can bound this bias. A lower degree of concentration will decrease the log price by exactly the amount that it increases log quantity. The variance of the log-price is therefore an upper bound for $Var(h)$: $Var(p_{1914}) = Var(h) + Var(\ln[\alpha + \gamma c]) \geq Var(h)$. Table 1 shows that among those dyes for which the 1914 import price is available (322 of the 508 imported), the variance of its log is .56, compared to 3.05 for that of log imports. Thus the variation in physical concentration biases the estimated coefficient on log imports downwards, and so biases $\hat{\beta}$ upwards, at most by a factor of 18 percent of the true coefficient value.

(d) Correcting for Non-Monopoly Behaviour among the Exporting Firms

Although we have assumed thus far that the demand for each dye is independent of all the others, the model is consistent with nested, non-address preferences, so long as those nests correspond to the included demand proxies. Let aggregate demand be representable by the utility function

$$U(q_1, q_2, ..., q_M) = G(\sum_g v_g (\sum_{i \in g} (\tilde{\alpha}_i \, q_i - (q_i / \tilde{S}_i)^{1+\gamma/\gamma})))$$

where g indexes the nests. Inverse demand is therefore

$$p_i = \frac{\partial}{\partial q_i} U(q_1, q_2, ..., q_M) = G'(.) v_g{}'(.) (\tilde{\alpha}_i - \frac{1+\gamma}{\gamma} (q_i / \tilde{S}_i)^{1/\gamma})$$

and Cournot residual demand is

$$D_i(p_i, q_{-i}) = S_i \mu_g (\alpha_i - p_i)^{\gamma}$$

where $S_i = \tilde{S}_i(\mu_g(1+\gamma)/\gamma)^{-\gamma}, \alpha_i = \tilde{\alpha}_i \, \mu_g$ and μ_g is a function of all the quantities in *i*'s nest, including that of *i* itself. If we assume, similar to Dixit and Stiglitz (1977), that firms do not take into account their effect on the sub-index function , then μ_g can be taken as exogenous to the firms, and we obtain our original model, with the product attributes proxying for differences in μ_g across the nests.

Finally, we allow for non-monopoly behaviour in exports to the U.S. in 1914 by assuming Cournot competition within each dye category as well. This alters the price and quantity and equations to

(5') $$P_i = [\alpha_i + \gamma N_i c_i]/(1 + \gamma N_i)$$

(6') $$Q_i = S_i [\frac{\gamma N_i}{1 + \gamma N_i} sign(\gamma)(\alpha_i - c_i)]^{\gamma}$$

(see Genesove and Mullin, 1998), so that one needs to adjust log imports by subtracting $\gamma \ln(1 + 1/\gamma N_i)$ from it in the expression for potential monopoly profits in equation (8):

(8') $$\pi_i = [q_{1914,i} + \gamma \ln(1 + \gamma^{-1} N_i^{-1})] + z_i + (\gamma + 1)(z_i^U - z_i)\gamma + e$$

Note that the adjusted log import is decreasing in N : given quantity, a larger number of firms implies a lower demand. It is increasing in $\kappa = \gamma / 1 + \gamma$, the monopoly marginal markup of cost. Also, c_i is now the (un-weighted) average marginal cost among the N firms.

Since γ is unknown, we consider a number of alternative values for it. For most of the analysis, we will assume $\gamma = 0$, and so $\kappa = 0$, for which no adjustment to imports is necessary. To recall, this is the case of completely inelastic demand up to a common reservation price. The opposite limiting case is that of $\gamma = -1$, corresponding to $\kappa = \infty$. But since the adjustment term is then incalculable when $N = 1$, we consider instead $\gamma = -1.1$, corresponding to an extremely large κ of ten. In addition, we consider linear ($\gamma = 1$, i.e., $\kappa = 1/2$) and exponential ($\gamma \to \pm\infty, \kappa = 1$) demand.

An alternative adjustment substitutes the inverse Hirschman-Herfindahl Index (HHI) for the number of firms. This is appropriate if the weighted (by market share) marginal cost of the foreign exporting firms provides a better comparison for U.S. costs in 1917 than does the un-weighted marginal cost. We consider this adjustment as well.

An alternative approach of treating γ as an estimable parameter would be appropriate, only if the number of firms were exogenous. But that will not be the case so long as z_i varies, since higher demand through higher α_i and lower costs would have induced more firms to have entered the dye category in the pre-war period. Furthermore, the resulting bias is indeterminate. To see this, consider the case of $|\gamma|$ large, so that the adjustment term is well approximated by $1/N$. Were the variable exogenous, we should then expect to estimate a coefficient on it equal to that on log imports. However, this variable will be negatively correlated with z_i, and negatively (positively) correlated with $(\gamma + 1)\{z_i^U - z_i)$, if γ is positive (negative).[22] Since $1/N$ is correlated with log imports, (negatively, as one would expect), this bias will be transferred to the coefficient on the latter, so that it is unclear whether including $1/N$ will reduce or exacerbate any bias generated by its absence. Nonetheless, we treat its inclusion as a robustness test, and consider regressions with it and without it.

Section V: American Production in 1917

---

[22] If e_i is at least in part forecastable, and if there are switching costs or the like, we would expect $1/N$ to be correlated with that error component as well.

Table 3 presents the linear probability regression of the incidence of US production in 1917. The sample is the 515 dyes that were imported in 1914, and for which a discovery year is available. The bivariate regression on the discovery year, shown in Column (1), estimates that a one year "older" dye is 1.7 percentage points more likely to have been produced. The corresponding scatter diagram, showing the fraction of dyes produced in 1917 by discovery year cohort, and with the regression line overlaid, is shown in Figure 2b.

Column (2) adds log imports. Its estimated coefficient implies that a dye whose imports were a hundred times greater than some other dye discovered in the same year was 29 percent (=.063*ln(100)) more likely to have been produced in 1917. As noted earlier, this coefficient has the interpretation of the probability of success ρ divided by the scale of the distribution of pre-development log expected profits, which two parameters can not be separately identified.[23] Including the variable has only a small effect on the coefficient on the year of discovery.

But it is $\hat{\beta}$, the ratio of the coefficient on the year of discovery to the coefficient on log imports, that is of central interest. This is -.25 and is reported in the third to last row of the table, with its standard error, calculated by the delta method, reported below.[24] The estimated ratio has the interpretation that a dye discovered one year later will cost 1-exp(.25) = 28 percent more to develop. This implies, in turn, that it was more than 13,000 times more expensive to develop the very latest dyes, from 1913, than the earliest ones in our sample, from 1875. We will have reason in the analysis that follows to revise these estimates downwards, but it is nonetheless worth noting at this point that these numbers, although large, are not completely unreasonable. It is not inconceivable that knowledge on the production of the 1875 cohort of dyes was nearly fully available, so that development costs for them would be little more than the cost of a simple run through. In contrast, the relevant knowledge for the 1913 cohort may have been completely absent, so that development of those dyes by the U.S. firms would have been akin to discovering the dyes from scratch. So it is instructive to

---

[23] In the deterministic model in which success is guaranteed by paying F_i, the inverse of the coefficient on log imports times $\sqrt{1/12}$ (the standard deviation of the standard uniform distribution) has the interpretation of the standard deviation of pre-development log expected profits. Here it is 4.6 (a factor about 100 in levels),

[24] Calculating the standard error from the corresponding nonlinear least squares model in which the coefficient on log imports is set to one, and the linear index divided by the scale parameter, yields very similar results.

compare the estimates here to Gambardella (1999), who cites rates of marketed to total synthesized compounds from 1 in 3000 to 1in 6000 in the closely allied pharmaceutical industry in the 1980s. Nonetheless, Beer (1959) cites a rate of marketed to total synthesized dyes of 37 in 2378, for Bayer in the late 1800s, which is at least an order of magnitude lower than those in Gambardella, and estimated here.

Column (3) adds patent indicators. Since recent dyes are more likely to be covered by a patent, and as the vast majority of US dye patents were held by the German dye firms, one might think that the discovery year merely proxies for patent protection. True, German patents were eventually confiscated, but in 1917 the legislation for that had not yet been passed.[25] "US patent in force" is a dummy indicator for a dye category developed after 1917 for which at least one US patent is listed in Schultz. Since a patent may indicate profitability, the count of countries with at least one issued patent in that dye category is also added. Both variables are insignificant, both singly and jointly (p-value of .20), indicating that, not withstanding the claims of the US firms at the time, which were to lead to the confiscation of the German owned patents, the foreign owned patents were apparently not an impediment to US development and production.[26] Recall that the count of countries with a patent had a positive and significant effect in predicting importation in 1914. Its insignificance in predicting 1917 production, conditional on log imports, suggests that log imports captures most of its effect on profitability.

An alternative explanation for the negative coefficient on discovery year is intrinsic differences in the difficulty of development, which we will term complexity. Column (4) adds a number of different proxies for this. The first is the number of intermediates used in production of the dye (from Shreve, 1922). Only 12 percent of dyes required a single intermediate, almost exactly half required two intermediates, a quarter required three and very few required more than four. Presumably the more intermediates used in production of a dye, the more difficult the development of the dye, for the simple reasons that (a) one (or one's domestic supplier) had to determine

---

[25] The Trading with the Enemy Act was passed in October, 1917, half a year after the U.S.'s entry into the war. That allowed the US government to seize most assets, but not patents. Seizing the latter was permitted only by an amendment to the act in November 1918, right before the armistice. Although once the U.S. entered the war in August 1917, protecting the patents might have been difficult for the German firms, much of the development decision reflected in the 1917 census would have preceded that date.

[26] Hesse (November 1915) also makes the point that only about a fifth of the dye categories were in any way covered by patents in the United States.

how to produce more intermediates, and (b) there were more chemical steps involved. Like all proxies for complexity, however, this may be expected to effect not only development but also production costs, although increases in either would make development and production less likely. Nonetheless, its estimated coefficient is insignificant. Likewise, indicators for the number and type of chemical reactions involved in the production of the dye also proved insignificant (not shown).

Column (4) adds the dummy variables for Schultz' sixteen dye classes. As noted earlier, the dyes in each class generally share a common tar-crude or even a basic intermediate. The set of class dummies are highly significant, with a p-value of .03 on the F-test. As Figure 6 shows, the estimated coefficient in Table 3 (relative to the omitted class, Azo dyes – the largest category) is positively related to that in Table 2. The Schultz number itself (again, normalized to lie between zero and one), is also added. It is highly significant, both statistically and economically, with larger Schultz numbers predicted a lower probability of production. One can only guess why the Schultz number is relevant to the analysis. As Figure 5 shows, the number that Schultz gave to each dye is strongly, although far from perfectly, correlated with the year of discovery. Essentially, the ordering is by class, but there was substantial time overlap among the classes. Perhaps it is an indicator of complexity: some dyes are derivatives of other dyes, and are ordered thus in Schultz; as the derivates can only be produced if those from which they are derived can be produced as well, derivative status is likely to negatively correlated with production. But the number of such dyes is few. Another possibility is that the Schultz number reflects the order in which Schultz and his earlier co-author Paul Julius became aware of the dye, and so provides additional information on the year in which knowledge on the dye was in the public domain beyond that given by the year of discovery. Yet another possibility is simply that Schultz ordered the dyes within each category by the degree of complexity. When higher order terms of the Schultz number are added, they are never significant.

The value for the joint F-test test on the dye classes, the number of intermediates and the Schultz number – the "technical" attributes – is 3.37, with a p-value of .0000. Note again that we can not tell whether these variables operate through the development cost (F) or the cost of production and so post-entry profitability. Controlling for these technological attributes has a clear effect on the coefficients of interest, decreasing the magnitude of both the discovery year and the log of imports, but of the former more, so that the ratio $\hat{\beta}$ decreases in magnitude to -.19. This implies that, controlling for the classes, a one year "younger" dye will cost 20 percent more to develop, and that the 1913 cohort will cost over 1,300 times than the 1875 cohort – a factor much more in line with the rates from Beer and Gambaredella given above.

Column (5) adds the demand attributes: the dummy variables for colour and materials. The F-tests show the colours to be only marginally significant, but the materials highly significant (overwhelmingly due to the negative effect of cotton –for reasons that are unclear). Together the demand attributes are significant in this column. But when in Column (6) we add the "technical" attributes back in, the demand attributes are jointly insignificant, with a p-value on the joint test of .37. Recalling that these variables were highly significant in predicting imports in 1914, with or without conditioning on the "technical attributes", this is consistent with log-imports reflecting most of the variance in demand.

Column (7) adds the 1914 market share of Swiss firms, the source of almost all non-German imports before the war. Its coefficient, which we might have expected to be negative, is clearly insignificant. This is at first surprising, given that the Swiss were obviously not blockaded. Swiss production and trade was clearly adversely affected by the war, as they relied on the Germans for dye intermediates, the resources for production of which were diverted to chemical warfare, and presumably Atlantic warfare might have made any importation difficult. But with Swiss dyes constituting almost two-thirds of the value of imports in 1915, it is clear that Germany was clearly adversely affected more, and so the insignificant result remains a mystery.

Table 4 repeats these regressions with the log of the import price added.[27] That cuts down the number of observations by a third, as this variable is not always reported in either Norton (1916) or Norton (November 1916). To recall, price is a perfect proxy for changes in physical composition, if the markup does not vary, and a very good proxy for the latter as well, so long as either the vertical shifter or the cost terms do not vary much. It is also an excellent proxy if the U.S. firms had no information on cost but needed to infer it from price. The estimated coefficient on log price is always negative, and highly significant so long as the technical attributes are not concluded. With the latter included, the magnitude of the coefficient falls by a third, and is barely

---

[27] More precisely, this is the log of the value of imports minus the log of the quantity of imports.

marginally significant. Both the negative sign and the reduced magnitude and significance upon the inclusion of the technical attributes is consistent with the log price proxying mainly for costs, which are expected to decrease the incidence of production both because of the higher marginal cost and so long as development costs and production costs are positively correlated. The estimate rate of knowledge, $\hat{\beta}$, is barely affected by the inclusion of the price.

Table 5 adjusts the log quantity of imports to account for the presence of more than one firm in1914, using the Cournot model, as described above. It presents the estimated ratio $\hat{\beta}$ and its standard error for the various values of γ, using either the inverse number of firms or the HHI, and including either only the year of discovery and the adjusted log quantity or all the controls considered in Column (6) of Table 3. We see that the estimated value is robust to the choice between HHI and the inverse number of firms, and across the first three values of γ. Extreme values of the marginal monopoly markup, as in the last column, lead to much greater absolute estimates of $\hat{\beta}$. The last column shows $\hat{\beta}$ when the log of the inverse number of firms, or HHI, is added as a regressor. Because, as argued earlier, this regressor is likely to be endogenous, these results need to be viewed with caution.

We can also relax the assumption that the change in horizontal demand shifter s_i from 1914 to 1917 (e_i) is uncorrelated with its 1914 value. It is easy to see that if, instead, it follows an AR(1) process, that $\hat{\beta}$ will be biased upwards. Thus if $S_{i,t} = S_{it-1}^{\rho} \exp(\delta_t v_{it})$, so that $s_{ti} = \delta_t + \rho s_{t-1,i} + v_{t,i}$, where t indexes the year, δ_t is a common calendar time shock and $v_{t,i}$ is an i.i.d. error term, then $\text{plim}\hat{\beta} = \beta / \rho^3$, absent all the other biases. We thus have been assuming that the predictable evolution in the demand for a dye, relative to that of other dyes, was sufficiently slow over time to take $\rho^3 \approx 1$. Is this indeed reasonable? s_{ti}, of course, is unobservable, and so its auto-regression can not be estimated. Using quantities in its place, for example by running 1917 log quantity on 1914 log imports, runs into the problem that, as noted, of the 508 dyes imported in 1914 and in our sample, only 123 were produced in 1917. Furthermore, quantities are reported for only 25. This additional censoring is due to the U.S. Tariff Commission's rule that quantity be reported only if there a sufficient number of firms producing the dye,[28] which leads to selection on the dependent variable. Indeed the effect of the selection is pretty obvious in a scatter plot of 1917 quantities against 1914 imports, shown in Figure 7. Observations below a hypothetical regression line extrapolated from the observations in the upper right corner of the graph seem to be missing in the lower left corner. Unfortunately, the usual selection correction methods are not of any use here given the extremely high censoring rate of 95%.[29]

An alternative approach is to gauge ρ from the relationship of current to lagged output for later years. Figure 8 presents such plots, while the top panel of Table 7 presents the corresponding OLS regressions. All the coefficients are estimated at less than one, although not significantly so in 1918 and 1923. The estimated coefficient in the pooled regression for all years (with standard errors clustered on the dye category) is 0.83, and significantly less than one. The selection problem is less serious here, with censoring rates of 9 to 22 percent. To deal with the remaining selection problem, the bottom panel drops those observations with a lagged quantity less than the medium lagged quantity for that year. Since dyes with large quantities in the previous year are much less likely to be "selected out", this in the spirit of Heckman's (1990) "identification at infinity" method. Now, only one post war year is significantly different from one. Although obviously this is in part due to the increased standard errors, the point estimate of ρ for three of the years is extremely close to one. Furthermore, the pooled estimate is now 0.9, and is also insignificantly different from one. Given that the observed demand attributes are also insignificant in Tables 3 and 4, one may conclude that relative demand for dyes was relatively stable over time. A less convinced reader might assume that dye firms deflated 1914 import quantities by about $0.9^3 = .72$ to predict prospective profits, but that would still yield a very high development cost-reducing knowledge diffusion rate.

Finally, we consider non-linearities in the effect of the discovery year. We assume a partial linear model in which the probability of U.S. production equals the sum of an unknown function of the discovery year and a linear index of all other variables. Following Yatchew (2003), we first obtain consistent estimates of the parameters of all the other independent variables by ordering the observations by

---

[28] Quantities are never reported for three or fewer firms, and always reported for five or more firms.

[29] Monte Carlo studies that investigate selection models never even consider censoring rates beyond 75% (Puhani, 2000).

discovery year, and running the first differenced regression,[30] excluding the discovery year, and then non-parametrically estimate the unknown function of the discovery year by non-parametric regression of the indicator variables minus the predicted linear term on the discovery year. The non-parametric component is shown in Figure 9, where we have used the lowess estimator, with a bandwidth of 19 years, which is half of the range of the data.[31] The estimates suggest that the knowledge diffuses only after a two decade lag (perhaps after some non-identifiable initial diffusion), and even after some 40 years in not completely diffused. Both the long diffusion period and the lag are surprising results. The long diffusion period seems at odd with the survey evidence, such as Mansfield (1985), whose respondents claim that half of all innovations leak out in a year or year and a half (and, specifically, the former, for chemicals), or Levin et al. (1987), where at least 87 percent of all innovations could be duplicated in less than five, and typically in less than three years. But such evidence is relevant to knowledge diffusion among competitors with a more or less equal knowledge bases, and not between incumbents and outside firms, as is the case here.

One tempting explanation for the lag is that knowledge diffusion really only begins after patent protection ends, at which time more firms begin producing the dye, thus suddenly increasing the possibility for knowledge diffusion. However, Table 2 shows that patent protection, whether current or at any time and in any country does not affect the likelihood of U.S. development of the dye.

Note that the non-paramteric result is foreshadowed by Figure 2b, which shows the bivariate relationship between the incidence of production and the discovery year. To check whether the non-linearity derives from the boundedness property of probabilities (for a number of the later cohorts, none of the dyes are produced), I estimated a number of different discrete choice models, including spline functions in the discovery year, with fixed knots that split the discovery year variable into segments of equal length. Table 7 presents the estimated $\hat{\beta}$ for each segment of the spline, and shows that the non-linearity. These show that both the non-linearity result and $\hat{\beta}$ under linearity are robust to the distributional assumption.

Section VI: Post-Entry Competition: 1923 Production & Imports

---

[30] Since there are 515 observations and only 38 different values for the discovery year, we get nearly the same results by running a fixed effects (for each discovery year) regression in this first stage.
[31] The coefficient on the log of imports is .053, with a standard error of .009.

This section considers whether post-entry competition between incumbents and new firms is also affected by the number of years since a dye's discovery. In particular, it considers the incidence of U.S. production in 1923, after the German firms were allowed back into the market. The discovery year is relevant if there are important learning by doing effects on costs, or other increasing dominance effects, so that the older the dye, the more efficient are the incumbents relative to the American entrants; or if the older the dye, the more knowledge on how to produce the dye cheaply (and now just how to produce it at all, at reasonable cost) has diffused. Recalling the earlier discussion, we can expect to measure the net effect only of knowledge diffusion and increasing dominance.

Since we will be considering only those dyes that were produced in the period before the Germans re-entered the market, which we take to be 1917-1919, we face a selection problem. Recalling equation (3), (and leaving aside both the missing terms identified in equation (8) and the stochastic nature of success in development), we expect a dye to have been produced in those years if $f_i - e_i \leq [e_0 - f_0] - \beta Y_i + \ln q_{1914,i}$. This will impart a bias to the estimates in a probability model of US production in 1923 on so long as $f_i - e_i$ is correlated with the error term in that regression. That will be so long as the demand shock in 1917-1919 is correlated with the demand shock in 1923, which must surely be the case.[32] However the bias will be small if most of the variance in the earlier production decision reflects development costs (f_i) and not demand (e_i).

The usual approach to handling this sort of selection problem is to employ a selection correction term, a la Heckman (1976). If we are not going to rely on functional form assumptions alone, we must have an excluded variable that predicts American production in 1917 but not in 1923. The obvious candidates here are time varying demand or cost determinants. There are three that suggest themselves: (a) indicators for uniform colours such as olive-green, khaki and navy blue that would have been in high demand in 1917-1919, but not thereafter, (b) indicators for the presence of non-German exporters in 1914, and (c) indicators for dyes that use intermediates which were heavily in demand for production of explosives during the

---

[32] More precisely, so long as the difference in post-entry profitability between 1917-1919 and 1914 is correlated with profitability in 1923. That will be so if, for example, there is persistence in demand shocks.

war. We have already seen that the first two sets of instruments do not predict production in 1917, however reasonable is the argument that they should, hwile the third set is co-linear with the class dummies.

As an alternative procedure, I restrict the sample further by including only those dyes that (a) were produced by American firms in at least one of 1917, 1918 and 1919, and (b) either produced or imported (but not both) in 1923. There are 95 such dyes.

To motivate this approach, consider a market environment where competition is sufficiently strong that at most one firm can operate in a dye category. Assume, further, that the most efficient firm operates, if its monopoly profits are positive. Write the difference in the log of variable monopoly profits and the log of (post-development) fixed costs, for the most efficient U.S. firm as $\pi_{1923}^{U} = X\gamma^{U} + \eta + \varepsilon^{U}$, and the same for the most efficient foreign firm as $\pi_{1923}^{D} = X\gamma^{D} + \eta + \varepsilon^{D}$. It is assumed that η, ε^{D} and ε^{U} are mutually independent unobserved shocks. The latter two should be understood as (transformations of) cost shocks, while η should be seen as a horizontal demand shifter. As such, it will be correlated with e (to recall, the horizontal demand shock in 1917, relative to 1914); it is this correlation which generates the selection bias. Note the implicit assumption, then, that variations in demand affect both domestic and foreign firms in equal proportions

The probability, conditional on X and η, that a US firm will be in the market, given that there is a firm in the market, is

$$\Pr\{I_U = 1 \mid I_U + I_D = 1, X, \eta\} = \Pr\{I_U = 1 \mid X, \eta\} / \Pr\{I_U + I_D = 1 \mid X, \eta\}$$

where I_D (I_D) equals one if a US (foreign) firm is in the market, and zero otherwise, and where we have used the assumption above that only one firm can operate profitably in a given category. If we now assume that the error terms ε^{U} and ε^{D} have independent, extreme-value distributions, we obtain

$$\Pr\{I_U = 1 \mid X, \eta\} = \exp(\eta + X\gamma^{U}) / \{1 + \exp(\eta + X\gamma^{U}) + \exp(\eta + X\gamma^{D})\}$$

and

$$\Pr\{I_U + I_D = 1 \mid X, \eta\} = [\exp(\eta + X\gamma^U) + \exp(\eta + X\gamma^D)]/\{1 + \exp(\eta + X\gamma^U) + \exp(\eta + X\gamma^D)\}$$

Thus

$$\Pr\{I_U = 1 \mid I_U + I_D = 1, X, \eta\} = \exp(X\gamma^U)/\{\exp(X\gamma^U) + \exp(X\gamma^D)\} = \{1 + \exp(X(\gamma^D - \gamma^U))\}^{-1}$$

Note that the common unobservable shock η has been eliminated, and so, under our assumptions, there is no selection bias. Note by the same logic, however, we would expect any common observable factor j for which $\gamma_j^D = \gamma_j^U$, such as a horizontal demand shifter, to be eliminated as well.

Table 8 shows the results.[33] (For the sake of comparison with the earlier tables, and without apology, it employs a linear probability model, instead of a logit.) Its major finding is that no matter what variables are included in the regression, the discovery year has no effect on the conditional probability of US production. The net effect of knowledge diffusion and increasing dominance effects after development is nil.

A second robust finding is that the quantity of log imports is insignificant as well. This is consistent with our assumption that most of the variation in that variable arises from the variance in the horizontal demand shifter s, which should affect foreign and domestic profits equally. Likewise, the number of countries with patents is also insignificant. However, the demand attributes are jointly significant.[34]

Note that the insignificance of the discovery year and log imports does not arise from the small sample. Although the standard errors are larger here than in the earlier tables (about twice as large, and so close to $\sqrt{515/95}$ - the square root of the ratio of

---

[33] Contrary to the model's assumption, there are dyes for which there were both imports and domestic production in 1923. The model can be easily extended to allow for that if I_U (I_D) is reinterpreted as indicating that the domestic (foreign) firm has the greater market share, and $I_U + I_D = 1$ is interpreted as indicating that there is either production or importation (or both). Unfortunately, domestic production is not always reported. However, if the sample is extended to include those dyes for which there is both reported domestic production and imports (of which there are 128), and with the variables thus redefined, we obtain similar results to Table 8.

[34] This is consistent with the above model and previous results if colour and material affect the vertical demand shifter (α). Because US and foreign firms have different costs, variations in α will affect log monopoly foreign and domestic profits differentially.

the number of observations in Tables 3 and 8), the variables would have remained significant even with those higher standard errors had the coefficient estimates remained the same. Instead the coefficient on the discovery year falls by an order of magnitude, and that on log imports by at least a third.

We see some indication that more complex dyes are more likely to be imported than produced by the American firms in that the estimated coefficient on the number of intermediates is always negative and significant; however, the coefficient on the Schultz number is now positive, and at times marginally significant. The technical variables are jointly significant. Also, when log price is added (thus reducing the sample to 75 observations), its estimated coefficient is negative, although only significant in the absence of the other variables. Nevertheless, the negative sign is consistent with the interpretation that costly dyes (and so those with higher 1914 import prices) were more complex and so relatively more difficult for the Americans to produce.

Section VII: Conclusion

The pattern of signs and significance of the variables used in the three central regressions in this study – the incidences of importation in 1914, U.S. production in 1917, and U.S. production in 1923, conditional on production or importation in that year and previous U.S. production - is summarized in Table 9. The technical attributes are significant in all three regressions, and suggest that the more complex dyes were less likely to be imported in 1914, produced domestically in 1917 and produced domestically in 1923. The demand attributes are significant in the first and last regression; they are jointly insignificant in predicting U.S. production in 1917, suggesting that the log of quantity imports is an adequate proxy for demand.

The discovery year behaves differently from either the demand or technical attributes, suggesting that it proxies for neither. It has no effect on importation in 1914, and no effect on U.S. production in 1923. Its only effect is on the probability of U.S. production in 1917: the older the dye, the more likely it was to have been produced. This pattern suggests that the discovery year reflects the diffusion over time from incumbents to potential, outside firms of the knowledge necessary to develop the dye. When the estimated coefficient of the discovery year in the 1917 production regression is divided through by the estimated coefficient on the log of imports, we obtain a rate of knowledge diffusion, where knowledge is measured in terms of the additional cost necessary for development. The estimated rate ranges from 19 percent to 27 percent a year, depending on the specification. Non-linear estimation implies that, except perhaps for an initial, non-identifiable component, knowledge did not diffuse to the outside firms for about two decades after the discovery, but diffused rapidly, though incompletely, after that.

REFERENCES

Arora and Rosenberg, “Chemicals: A U.S. Success Story,” in Ashish Arora, Ralph Landau and Nathan Rosenberg, eds., *Chemicals and Long-Term Economic Growth: Insights from the Chemical Industry*, John Wiley & Sons, 1998.

Arora, A. “Appropriating Rent from Innovation: A Historical Look at the Chemical Industry” in Horst Albach and Stephanie Rosenkranz (eds.), *Intellectual Property Rights and Global Competition*, Verlag, Berlin, 1995a.

______, “Licensing Tacit Knowledge: Intellectual Property Rights and the Market for Know-How*”, Economics of Innovation and New Technology*, 4:41-49, 1995b.

Athey, S. and Schmutzler, A. “Investment and Market Dominance”, RAND Journal of Economics, 32(1):1-26, 2001.

Beer, J. *The Emergence of the Dye Industry*, 1959, Illinois.

Bresnahan, Timothy F., Reiss, Peter C. , “Entry in Monopoly Markets”, *Review of Economic Studies*. October 1990; 57(4): 531-53

Bresnahan, Timothy F., Reiss, Peter C. , “Entry and Competition in Concentrated Markets”, *Journal of Political Economy*. October 1991; 99(5): 977-1009.

Cabral, Luis and Michael Riordan, “The Learning Curve, Market Dominance, and Predatory Pricing”, *Econometrica*, 62:1115-1140, 1994.

Cantwell, John. “Technological lock-in of large firms since the interwar period”, *European Review of Economic History*, 4:147-174, 2000.

Cantwell, John, “The Evolution of European Industrial Technology in the Interwar Period” in Francois Caron, Paul Erker and Wolfram Fischer, eds., *Innovations in the European Economy between the Wars*, Walter de Gruyter, Berlin, New York, 1995.

Carleton, P. W. *Prices of Natural Dyestuffs and Tanning chemicals*, War Industries Board Price Bulletin, No. 52, 1919.

Dixit, Avinash and Joseph Stiglitz, “Monopolistic Competition and Optimum Product Diversity”, *American Economic Review*. 67(3): 297-308, 1977.

Doyle, Aida M. *Digest of Patents Relating to Coal-Tar Dyes and Allied Compounds: United States patents issued prior to January 1, 1924*, The Chemical Publishing Company, Easton, Pennsylvania, 1926.

Dyestuff Hearings before the Committee on Ways and Means House of Representatives on H.R. 2706 A Bill to Amend an Act Entitled “An Act to Increase the Revenue, *and for Other Purposes," Approved September 8, 1916, and H..R. 6495 A Bill to Regulate the Importation of Coal-Tar Products, to Promote the Establishment of the Manufacture thereof in the United States, and, as Incident thereto, to Amend the Act of September 8, 1916, entitled, "An Act to Increase the Revenue, and for Other Purposes"* June 18-20 and July 14-18, 1919, Government Printing Office, Washington, 1919.

Epstein, Stephan R., "Property Rights to Technical Knowledge in Premodern Europe, 1300-1800, *Papers and Proceedings*, 94(2), 382-387.

Gambardella, Alfonso, *Science and Innovation: The US pharmaceutical industry during the 1980s*, Cambridge University Press, 1999.

Genesove, David and Mullin, Wallace P., 'Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914", *RAND Journal of Economics*. Summer 29(2): 355-, 1998.

Gilbert, Richard and David Newbery, "Preemptive Patenting and the Persistence of Monopoly", American Economic Review, 72:514-526, 1982

Haber, L. F., *The Chemical Industry, 1900-1930*, Clarendon Press, Oxford, 1971.

Haynes, *.American Chemical Industry: A History*, D. van Nostrand Company, Inc.

Hearings Before the Committee on Ways and Means House of Representatives on H.R. 2706, A bill to Amend an act entitled "An Act to Increase the Revenue, and for Other Purposes", Approved September 8, 1916 and H.R. 6495, A bill to Regulate the Importation of Coal-Tar Products to Promote the Establishment of the Manufacture Thereof in the United States and, as incident thereto, to amend the Act of September 8, 1916, Entitled "An Act to increase the revenue and for other purposes", 1919, Washington, Government Printing Office.

Heckman, J.J., "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, " *Annals of Economic and Social Measurement*, 5, 475-492,1976.

_____, 1990.

Hesse, Bernhard C., "Lest We Forget! Who Killed Cock Robin? The U.S. Tariff-History of Coal-Tar Dyes", " *The Journal of Industrial and Engineering Chemistry*, August 1915, 694-709.

_____, "Coal-Tar Dyes and the Paige Bill", *The Journal of Industrial and Engineering Chemistry*, November 1915, 963-974.

Hounshell and Smith, *Science and Corporate Strategy, Du Pont R&D, 1902-1980*. Cambridge University Press, 1988.

Hohenberg, Paul M. *Chemicals in Western Europe: 1850-1914: An Economic Study of Technical Change*, Read McNally and Company, Chicago, 1967.

Jones, W. N. and F. W. Cassebeer, *Prices of Coal-Tar Crudes, Intermediates and Dyes*, War Industries Board Price Bulletin No. 53, Washington, Government Printing Office, 1919.

Klepper, S., "Entry, Exit, Growth and Innovation over the Product Life Cycle", *American Economic Review*, 86:562-583, 1996.

Kudo, Akira, "I.G. Farben in Japan: The Transfer of Technology and Managerial Skills" in *Business History*, 36(1):159-183,1994.

Liebenau, Jonathan, "The Management of High Technology: The Use of Informationin the German Chemical Industry, 1890-1930" in A. Kudo and T. Hara, eds., *International Cartels in Business History*, University of Tokyo Press, Tokyo, 57-73, 1992.

Levin, R.C., Klevorick, A.K., Nelson R.R. and Winter, S. G. "Appropriating the returns from industrial R&D", *Brookings Papers on Economic Activity*, 783-820, 1987.

Little, Arthur D., "The Dyestuff Situation and Its Lesson", *The Journal of Industrial and Engineering Chemistry*, March 1915, 237-239.

Mansfield, E., "How Rapidly Does New Industrial Technology Leak Out"" *Journal of Industrial Economics*, 34 (December 1985) 217-223.

Marsch, Ulrich, "Strategies for Success: Research Organization in German Chemical Companies and IG Farben Until 1936", *History and Technology*, 12:23-77, 1994.

Murmann, Johann Peter, *Knowledge and Competitive Advantage: The Co-evolution of Firms, Technology, and National Institutions*, Cambridge University Press, 2003.

Murmann, Johann Peter and Ralph Landau, "On the Making of Competitive Advantage: The Development of the Chemical Industries of Britain and Germany Since 1850" in Ashish Arora, Ralph Landau and Nathan Rosenberg, eds., *Chemicals and Long-Term Economic Growth: Insights from the Chemical Industry*, John Wiley & Sons, 1998.

Norton, Thomas H., *Artificial Dyestuffs Used in the United States: Quantity and Value of Foreign Imports and of Domestic Production during the Fiscal Year 1913-1914,*Government Printing Office, Washington, 1916.

_______, "A Census of the Artificial Dyestuffs Used in the United States" *The Journal of Industrial and Engineering Chemistry*, November 1916, pages 1039-1048.

_______, *Dyestuff Situation in the United States November 1915*, Reprinted and Revised May 1, 1916, Government Printing Office, 1916.

_______, "Dyestuff Situation in the United States at the Close of 1915", *The Journal of Industrial and Engineering Chemistry*, February 1916, pages 166-172.

_______, "U.S. Dyestuff Situation at Close of 1915 – Correction", *The Journal of Industrial and Engineering Chemistry*, March 1916, page 287.

_______, *Artificial Dyestuffs Imported into the United States 1913-14, Supplemental Statistics to Accompany Special Agents Series No. 121*, Department of Commerce, Bureau of Foreign and Domestic Commerce, 1922.

Lawrence H. Officer, "Exchange rate between the United States dollar and forty other countries, 1913 -1999." *Economic History Services*, EH.Net, 2002. URL: http://www.eh.net/hmit/exchangerates/

Puhani, P.A., "The Heckman Correction for Sample Selection and Its Critique", *Journal of Economic Surveys*, 14(1):53-68(16) ,2000.

Putnam, Jonathan., "The Value of International Patent Rights, mimeo,1997.

Reader, W. J., *Imperial Chemical Industries: A History, Volume I: The Forerunners 1870-1926*, Oxford University Press, London, 1970.

Rowe, F.M., ed., *Colour Index*, .Society of Dyers and Colourists 1924.

Rowe, F.M. ed., *Supplement to the Colour Index*: The Society of Dyers and Colourists, 1928.

Schröter, Verena, "Participation in market control through foreign investment: IG Farbenindustrie AG in the United States: 1920-38" in Alice Teichova, Maurice Lévy-Leboyer and Helga Nussbaum, *Multinational enterprise in historical perspective*, Cambridge University Press, 1986.

Schultz, Gustav. *Farbstofftabellen*, (5th edition, anastatischer Neudruck, 1920), Berlin, Weidmannsche Buchhandlung 1914.

Shreve, R. Norris, (in collaboration with Warren N. Watson and A. R. Willis), *Dyes Classified by Intermediates*, The Chemical Catalog Company, New York, 1922.

Steen, Kathryn, *Wartime Catalyst and Postwar Reaction: The Making of the United States Synthetic Organic Chemicals Industry, 1910-1930*, Dissertation, University of Delaware, August 1995a.

_______, "Confiscated Commerce: American Importers of German Synthetic Organic Chemicals, 1914-1929", *History and Technology,* 12:261-284, 1995b.

_______, "Patents, Patriotism, and "Skilled in the Art": USA v. The chemical Foundation, Inc., 1923-1926", *Isis*, 92:91-122, 2001.

Taussig, Frank, "The Tariff Act of 1922", *Quarterly Journal of Economics*, 37(1):1-28, 1922.

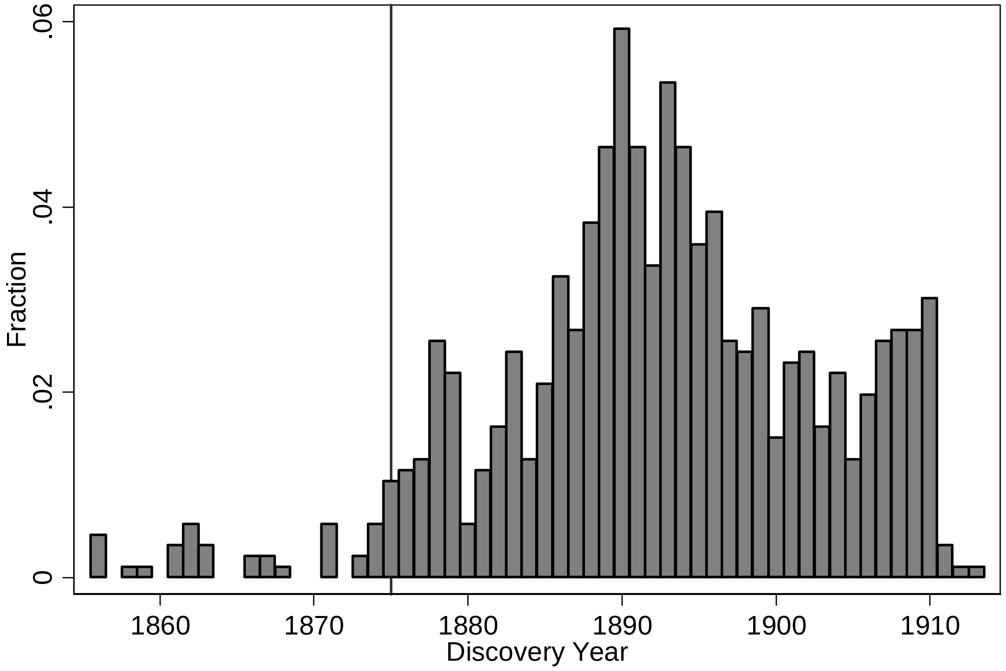
Travis, Anthony, *Dyes Made in America, 1915-1980, The Calco Chemical Company, American Cyanamid and the Raritan River,* Sidney M. Edelstein Center for the History and Philosophy of Science, Technology and Medicine at the Hebrew University of Jerusalem and the Hexagon Press, 2004.

United States Tariff Commission, *Census of Dyes and Coal-Tar Chemicals*, 1917.

United States Tariff Commission, *The Dyestuff Situation in the Textile Industries*, 1918, Washington

Yarchew, Ardonis, *Semiparametric Regression for the Applied Econometrician*, Cambridge University Press, Cambridge, 2003.

Figure 1. The Distribution of Dyes by Year of Discovery



Year of discovery is as recorded in Schultz (1914). The sample used in the paper is for dyes discovered between 1875 and 1913, inclusive.

Figure 2a. Fraction of Dyes Imported in 1914
by Discovery Year

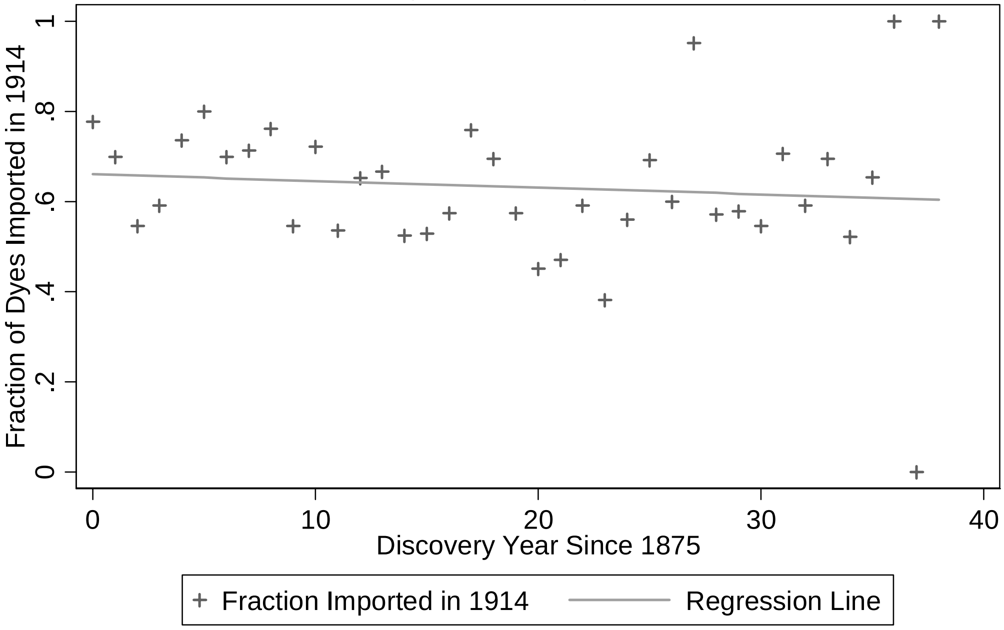
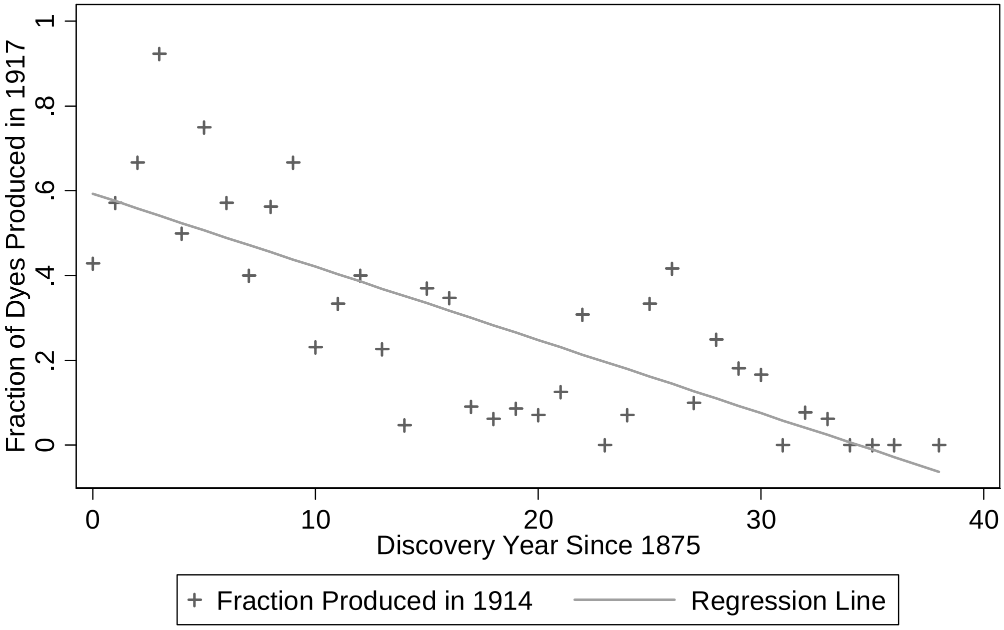


Figure 2b. Fraction of Dyes Produced in 1917
(of dyes imported in 1914) by Discovery Year



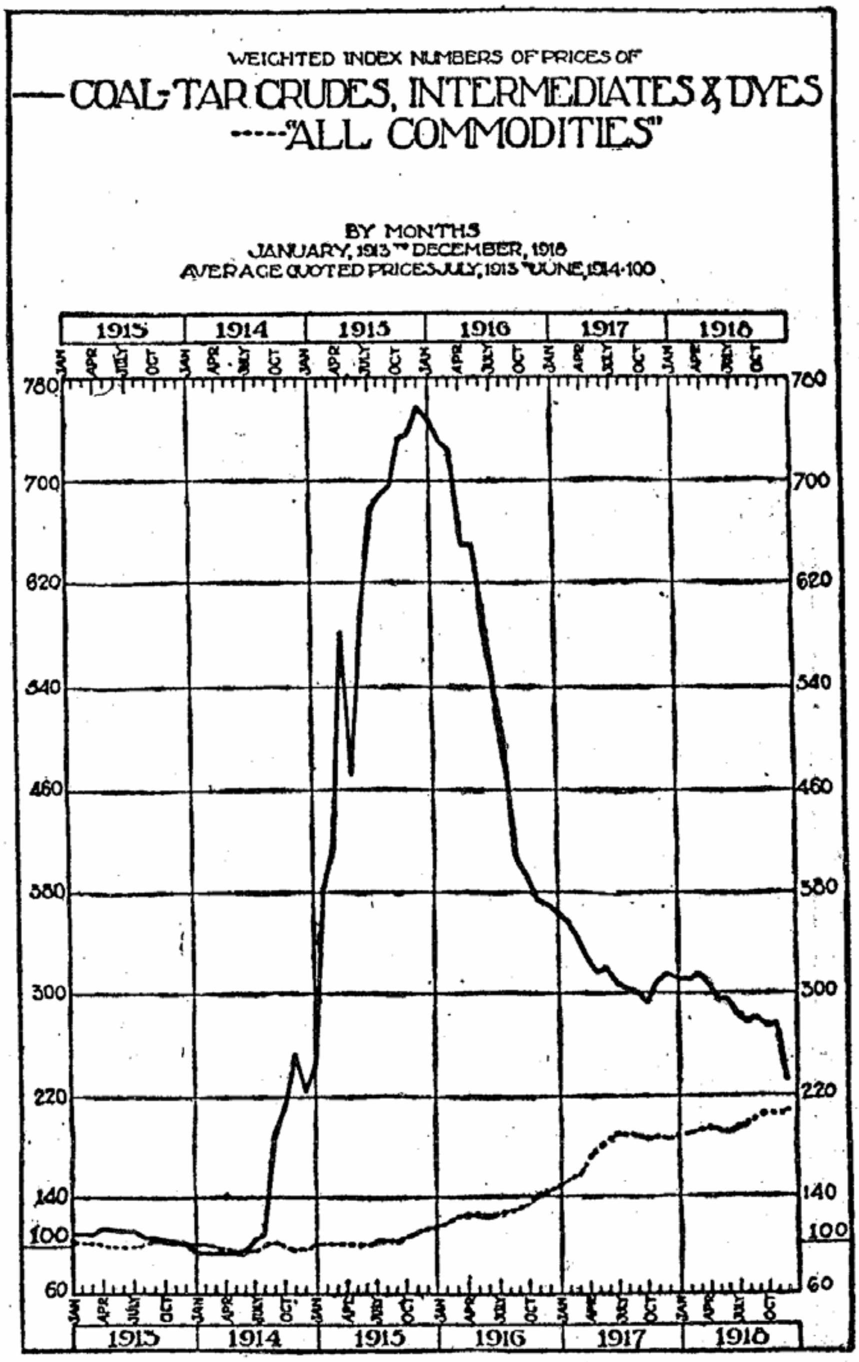

WEIGHTED INDEX NUMBERS OF PRICES OF
— COAL-TAR CRUDES, INTERMEDIATES & DYES
----- "ALL COMMODITIES"
BY MONTHS
JANUARY, 1913 TO DECEMBER, 1918
AVERAGE QUOTED PRICES JULY, 1913 TO JUNE, 1914 = 100
1913
1914
1915
1916
1917
1918
JAN
APR
JULY
OCT
780
700
620
540
460
380
300
220
140
100
60

FIGURE 3

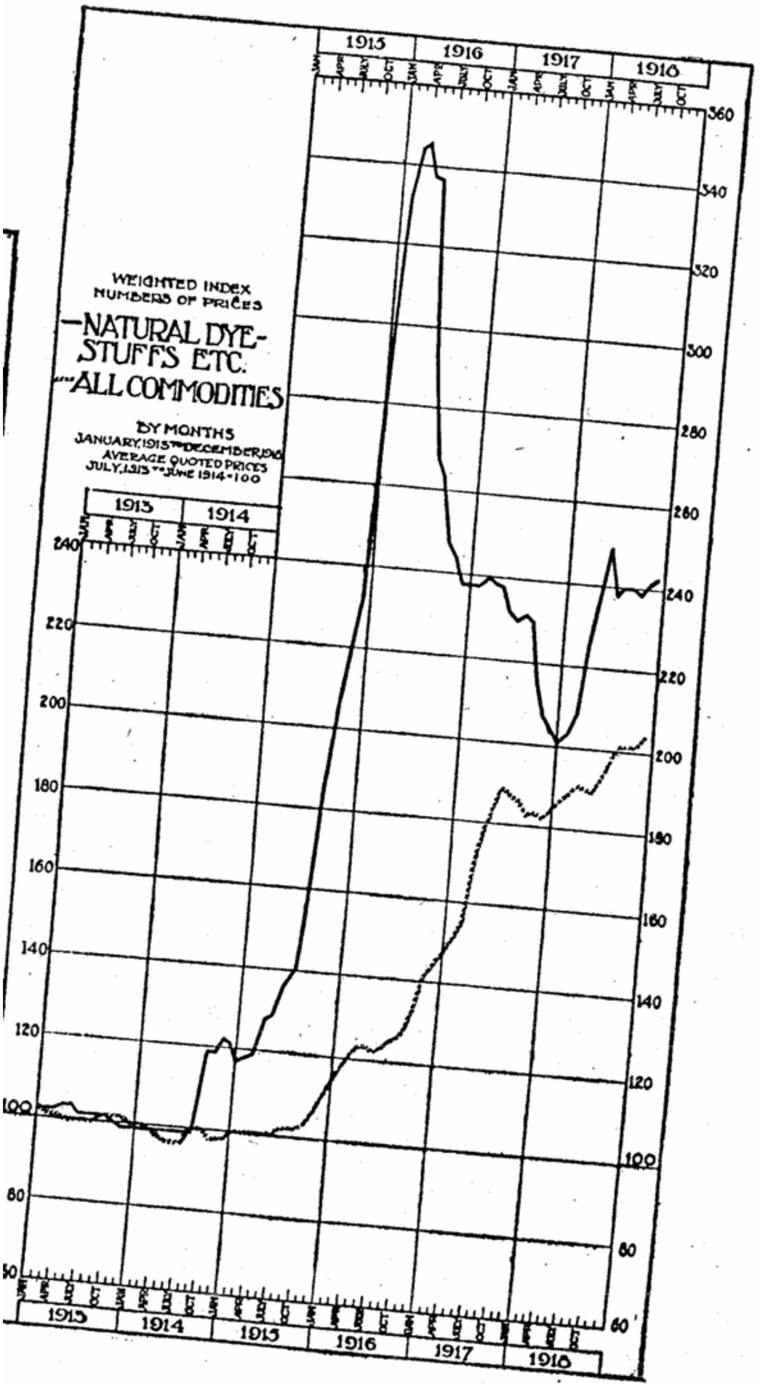

WEIGHTED INDEX NUMBERS OF PRICES
—NATURAL DYE-STUFFS ETC.
.....ALL COMMODITIES
BY MONTHS
JANUARY, 1913 TO DECEMBER 1918
AVERAGE QUOTED PRICES
JULY, 1913 TO JUNE 1914=100
1913
1914
1915
1916
1917
1918
JAN
APR
JULY
OCT
360
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60

FIGURE 4

Figure 5: Schultz Number and Discovery Year

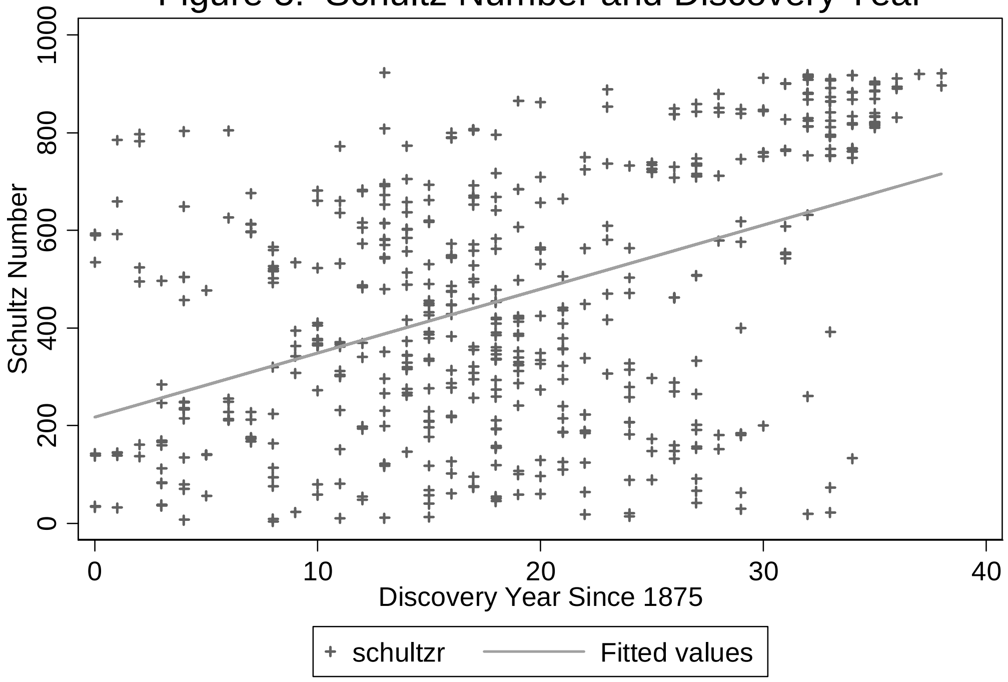
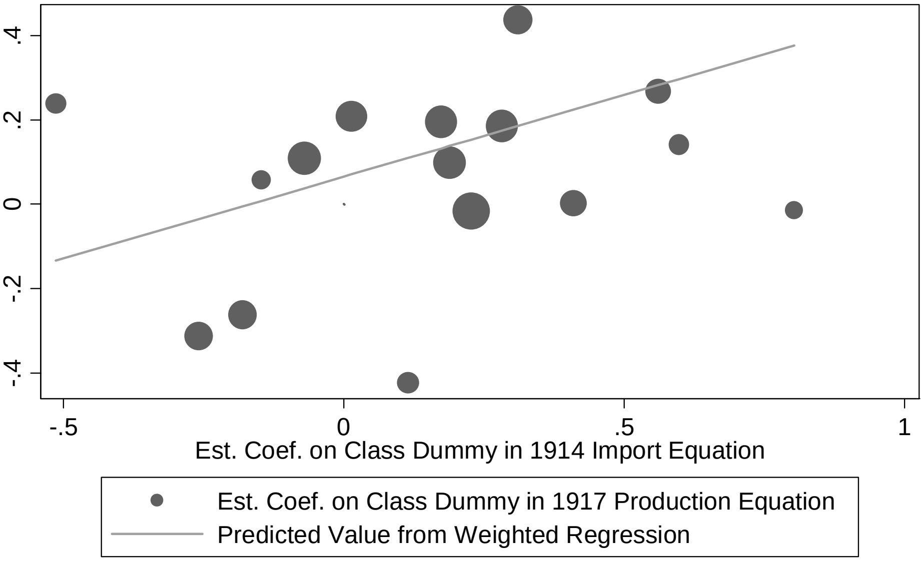


Figure 6: Regression Coefficients from 1914 Imports and 1917 Production



Reg Line: .06 + .40*1914 Coef. Estimate
(.35) (.21)

Figure 7

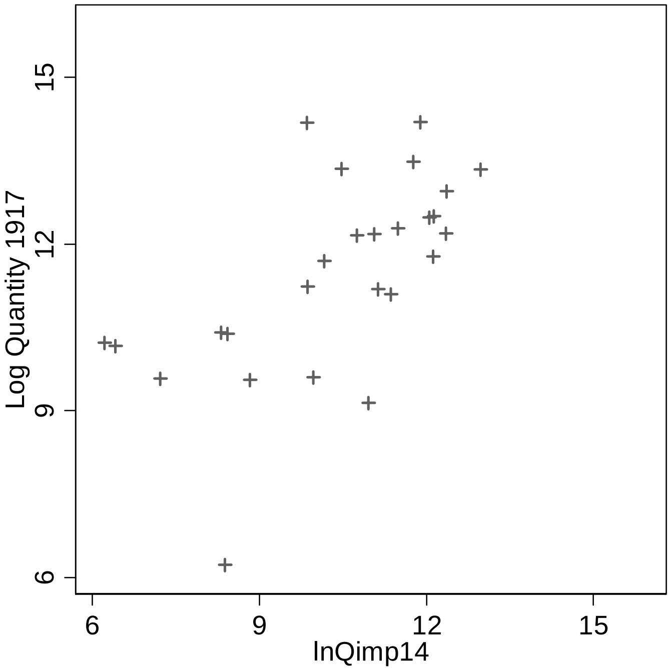


Figure 8: Year to Year Production Quantities

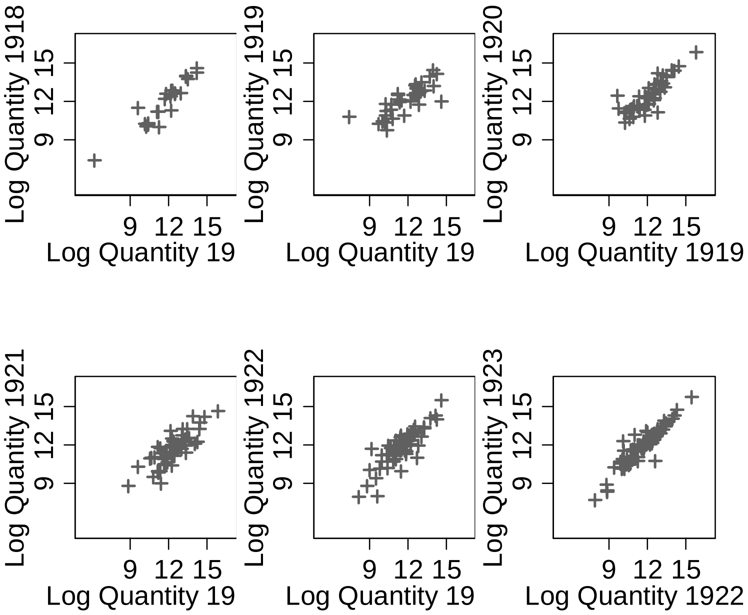
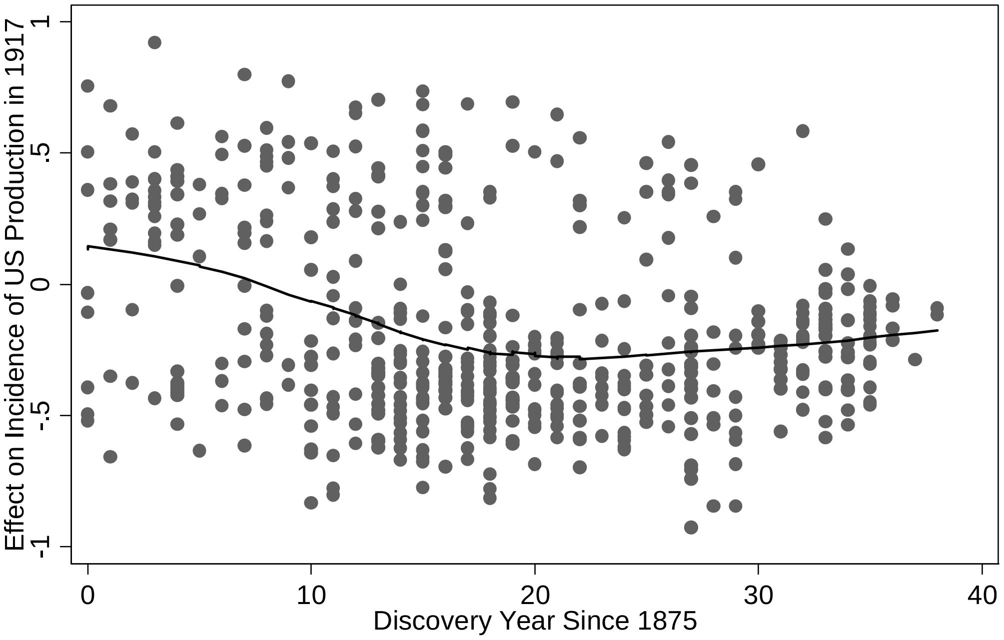


Figure 9 Non-parametric Estimate in Partial Linear Model



bandwidth = .5

Table 0: Importation in 1914 and American Production in 1917

	Not Produced in 1917	Produced in 1917	All
Not Imported in 1914	300	20	320
Imported in 1914	390	125	515
All	690	145	837

Table 1: Summary Statistics

	All post 1875 Dyes (N=835)		Imported in 1914 (N=515)		Imported in 1914 & Price reported (N=345)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Discovery Year - 1875	18.67	9.24	18.53	9.57	17.39	9.62
Ln (Import Quant)			8.85	2.09	9.74	1.74
Import Quantity (tonnes)			39.2	93. 0	56.3	108.6
Ln Import Price					-1.42	0.73
Price (dollars)					0.32	0.33
Schultz Number	0.50	0.29	0.50	0.29	0.47	0.29
No. of Intermediates	2.36	1.01	2.36	1.03	2.37	1.00
	Number	%	Number	%	Number	%
MATERIALS						
Cotton	501	*61.0*	304	*60.2*	185	*57.8*
Silk	188	*22.9*	138	*27.3*	105	*32.8*
Wool	391	*47.6*	254	*50.3*	183	*57.2*
COLOURS:						
Black	89	*10.8*	48	*9.5*	30	*9.3*
Blue	161	*19.5*	103	*20.2*	66	*20.5*
Red	76	*9.2*	57	*11.2*	34	*10.6*
Yellow	72	*8.7*	47	*9.3*	35	*10.9*
Pink	5	*0.6*	4	*0.8*	1	*0.3*
Purple	4	*0.4*	2	*0.4*	1	*0.3*
Brown	62	*7.5*	41	*8.1*	25	*7.8*
Green	50	*6.1*	36	*7.1*	21	*6.5*
Orange	45	*5.4*	39	*7.7*	22	*6.8*
Violet	60	*7.3*	45	*8.9*	28	*8.7*
Gold	4	*0.5*	3	*0.6*	2	*0.6*
Grey	9	*0.1*	7	*1.4*	4	*1.2*
Scarlet	31	*1.1*	22	*4.3*	17	*5.3*

Olive	5	*3.8*	4	*0.8*	2	*0.6*
CLASS:						
Nitroso	4	*0.5*	1	*0.2*	1	*0.3*
Nitro	2	*0.2*	1	*0.2*	1	*0.3*
Stillbene	10	*1.2*	6	*1.2*	5	*1.6*
Pyrazolone	12	*1.5*	5	*1.0*	3	*1.6*
Azo (omitted dum)	436	*52.7*	272	*53.5*	173	*54.0*
Auramine	2	*0.2*	2	*0.4*	2	*0.6*
Triphenyl	55	*6.7*	43	*8.5*	34	*10.6*
Xanthone	25	*3.0*	15	*3.0*	11	*3.4*
Acridine	7	*0.9*	6	*1.2*	5	*1.6*
Quinoline	3	*0.4*	2	*0.4*	2	*0.6*
Thiobene	5	*0.6*	5	*1.0*	4	*1.2*
Indophen	1	*0.1*	0	*0.0*	0	*0.0*
Oxazine	46	*5.6*	17	*3.4*	10	*3.1*
Azine	31	*3.8*	14	*3.8*	5	*1.6*
Sulphur	48	*5.8*	27	*5.3*	15	*4.7*
Anthraq	108	*13.1*	70	*13.8*	32	*9.9*
Indigo	32	*3.9*	22	*4.3*	17	*5.3*

Table 2: Incidence of American Importation in 1914

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	10)
Discovery Year – 1875	-.001	-.001	-.002	-.001	-.001	-.002	-.0004	-.002	-.001	-.002
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
# Countries w/ Patents		.019								.024
		(.011)								(.011)
16 Dye Classes			2.36					2.63		2.23
(p-value)			(.002)					(.0005)		(.004)
Schultz Number				-.008				-.34		-.44
(Norm. on [0,1])				(.064)				(.17)		(.19)
No. of Intermediates					-.002			.01		.019
					(.017)			(.02)		(.019)
16 Colours						3.34			3.27	3.18
(p-value)						(.0000)			(.0000)	(.0000)
3 Materials							4.49		4.16	2.50
(p-value)							(.004)		(.006)	(.058)
Constant	.64	.60	.66	.64	.64	.53	.55	.72	.44	.47
	(.04)	(.05)	(.04)	(.04)	(.06)	(.04)	(.06)	(.07)	(.06)	(.09)
Technical Attributes								2.34		2.01
(p-value)								(.013)		(.008)
Demand Attributes									3.51	3.17
(p-value)									(.0000)	(.0000)
R-squared	.0004	.004	.045	.0004	.0004	.058	.016	.049	.072	.121
No. of Observations.	835	835	835	835	835	835	835	835	835	835

Where a set of dummy variables are used, the F-test, and below it the corresponding p-value, is reported. Otherwise, standard errors are reported within the parentheses. Note that for colours and materials more than one dummy can equal one, as a dye can produce more than one colour and be used on more than one type of material.

Table 3: Incidence of American Production in 1917

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Discovery Year – 1875	-.017	-.016	-.016	-.012	-.014	-.013	-.013
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Ln(Import Quantity)		.063	.062	.066	.063	.066	.065
		(.008)	(.008)	(.008)	(.008)	(.008)	(.008)
No. Countries w/ Patents			.004				.012
			(.017)				(.011)
U.S. Patent in Force			.008				
			(.050)				
Number of Intermediates				-.013		-.007	-.007
				(.017)		(.018)	(.018)
16 Dye Classes				1.83		1.72	1.68
(p-value)				(.03)		(.04)	(.05)
Schultz Number				-.77		-.57	-.57
				(.16)		(.20)	(.20)
14 Colours					1.41	.96	.94
(p-value)					(.14)	(.50)	(.52)
3 Materials					5.98	2.15	2.14
(p-value)					(.0005)	(.09)	(.09)
Technical Attributes				*3.22*		*2.13*	*2.11*
(p-value)				*(.0000)*		*(.005)*	*(.005)*
Demand Attributes					*2.17*	*1.11*	*1.10*
(p-value)					*(.004)*	*(.34)*	*(.35)*
Swiss Market Share							-.025
							(.050)
R-squared	.15	.24	.24	.32	.30	.35	.35
Ratio of Coefficients $(\hat{\beta})$	---	-.25	-.25	-.19	-.22	-.20	-.20
	(---)	(.04)	(.05)	(.04)	(.04)	(.05)	(.05)
Number of Observations	515	515	515	515	515	515	515

The sample is all dyes imported in 1914 with a reported discovery year of 1875 or later. Ratio of Coefficients $(\hat{\beta})$ is the ratio of the coefficient on discovery year to the ratio of the coefficient on Ln(Import Quantity). Its standard error is calculated by the delta method, as implemented in Stata's nlcom command. The F-test is reported for set of dummy variables, with its p-value in parentheses below. Otherwise, standard errors are reported within the parentheses. Joint test for more than one set of variables, or a set and one or more variables are shown in italics.

Table 4: Incidence of American Production in 1917 (with log-Price)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Discovery Year – 1875	-.020	-.018	-.018	-.014	-.017	-.016	-.016
	(.002)	(.002)	(.002)	(.003)	(.003)	(.003)	(.003)
Ln(Import Quantity)		.065	.065	.075	.066	.075	.074
		(.013)	(.013)	(.013)	(.014)	(.015)	(.015)
Ln(Import Price)		-.074	-.073	-.052	-.066	-.054	-.053
		(.032)	(.031)	(.036)	(.035)	(.038)	(.038)
No. Countries w/ Patents			.007			.004	
			(.021)			(.015)	
U.S. Patent in Force			-.028				
			(.065)				
Number of Intermediates				-.007		-.006	-.006
				(.023)		(.025)	(.025)
16 Dye Classes				1.91		1.91	1.88
(p-value)				(.02)		(.02)	(.02)
Schultz Number				-.80		-.70	-.70
				(.02)		(.27)	(.27)
14 Colours					.76	.80	.81
(p-value)					(.72)	(.68)	(.68)
3 Materials					2.52	.90	.90
(p-value)					(.06)	(.44)	(.44)
Technical Attributes				*2.35*		*2.03*	*2.02*
(p-value)				*(.002)*		*(.009)*	*(.009)*
Demand Attributes					*1.05*	*.81*	*.81*
(p-value)					*(.41)*	*(.68)*	*(.69)*
Swiss Market Share							-.017
							(.07)
R-squared	.17	.26	.26	.35	.30	.38	.38
Ratio of Coefficients ($\hat{\beta}$)	---	-.27	-.27	-.19	-.25	-.21	-.21
	(---)	(.07)	(.07)	(.05)	(.07)	(.06)	(.06)
Number of Observations	345	345	345	345	345	345	345

The F-test is reported for set of dummy variables, with its p-value in parentheses below. Otherwise, standard errors are reported within the parentheses. Joint test for more than one set of variables, or a set and one or more variables are shown in italics.

Table 5: Adjusting Quantity Imports for the (Equivalent) Number of Firms

		(1)	(2)	(3)	(4)	(5)
Control Variables from Column (6) of Table 3 Included?	1/N or HHI	$\gamma = 0$	$\gamma = 1$	$\gamma = \infty$	$\gamma = -1.1$	Endogenous Inverse No. of Firms (or HHI)
No	1/N	-.25	-.26	-.27	-.37	-.27
		(.04)	(.06)	(.05)	(.07)	(.07)
Yes	1/N	-.20	-.20	-.21	-.29	-.19
		(.04)	(.05)	(.05)	(.06)	(.06)
No	HHI	as above	-.26	-.27	-.36	-.26
			(.05)	(.05)	(.07)	(.07)
Yes	HHI		-.21	-.21	-.29	-.19
			(.05)	(.05)	(.07)	(.07)

Table 6: American Production in 1917 (with non-linearities)

	(1) No Spline	(2) Two Splines	(3)	(4) F-Test for Equality of Splines (p-value)
Uniform (Linear Prob.)	-.20	-.46	.13	22.33
	(.05)	(.10)	(.08)	(.0000)
Normal (Probit)	-.20	-.33	.04	7.49
	(.05)	(.08)	(.10)	(.006)
Logistic (Logit)	-.20	-.32	.03	6.63
	(.05)	(.08)	(.10)	(.010)
Log Log	-,21	-.37	.07	9.51
	(.05)	(.09)	(.10)	(.002)
C Log Log	-.20	-.29	-.005	4.73
	(.05)	(.07)	(.09)	(.030)

Control variables from column (6) of Table 3 are included

Table 7: Auto-regressions of Log Quantity

	(1) 1917	(2) 1918	(3) 1919	(4) 1920	(5) 1921	(6) 1922	(7) 1923	(8) 1918-23
Panel A:								
Log Lagged Quantity	0.50	0.99	0.59	0.81	0.77	0.87	0.95	0.83
	(.13)	(.11)	(.08)	(.08)	(.09)	(.07)	(.05)	(.05)
Constant	6.44	.31	5.16	2.63	2.08	1.82	0.88	Year
	(1.35)	(1.27)	(.94)	(.92)	(1.07)	(.85)	(.56)	dummies
R-squared	25	.82	.62	.73	.64	.71	.85	.73
Percent Censored	95%	10%	19%	18%	22%	9%	19%	
Number of Observations	25	22	36	46	49	60	70	283
Panel B:								
Log Lagged Quantity	1.31	1.12	0.43	1.01	0.77	0.97	0.98	0.90
	(.50)	(.19)	(.23)	(.15)	(.15)	(.15)	(.09)	(.07)
Constant	-3.33	-1.31	7.39	0.09	2.13	0.45	0.50	Year
	(5.95)	(2.43)	(3.05)	(1.95)	(1.96)	(1.89)	(1.20)	dummies
R-squared	.38	.78	.16	.68	.55	.59	.76	.63
Number of Observations	13	12	19	24	25	31	36	147

Lagged Quantity is that of the previous year, except for Column (1) (1917), where the quantity of imports in 1914 is used. The sample in each column consists of those dye categories with a recorded quantity in the (i) current and (ii) previous year, (iii) a recorded discovery year and (iv) importation in 1914. Percent observed is the number of observations divided by the number of observations in the set defined by conditions (ii), (iii) and (iv) only.

Table 8: US Production in 1923, among dyes that were (a) produced in 1917-1919 & (b) produced or imported in 1923

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Discovery Year – 1875	-.002	-.002	-.0003	-.001	.002	.0001	-.003	.005
	(.004)	(.004)	(.004)	(.004)	(.005)	(.005)	(.005)	(.006)
Ln(Import Quantity)		.019	.014	.022	.007	-.003	.004	.005
		(.016)	(.016)	(.016)	(.018)	(.018)	(.022)	(.024)
Ln(Import Price)							-.107	-.138
							(.060)	(.086)
Countries w/ Patents			.024			.028		.028
			(.023)			(.025)		(.031)
# of Intermediates				-.11		-.13		-.15
				(.05)		(.06)		(.06)
Schultz Number				.29		.63		.75
				(.28)		(.36)		(.42)
12 Dye Classes				2.43		3.00		2.14
(p-values)				(.010)		(.003)		(.04)
12 Colours					1.25	2.04		1.21
(p-values)					(.27)	(.035)		(.31)
3 Materials					2.30	3.94		2.67
(p-value)					(.08)	(.01)		(.06)
Technical Attributes				2.62		2.54		1.96
(p-value)				(.004)		(.007)		(.05)
Demand Attributes					1.81	2.44		1.76
(p-value)					(.05)	(.007)		(.08)
R-squared	.002	.016	.03	.33	.27	.53	.08	.60
Number of Observations	95	95	95	95	95	95	75	75

The F-test is reported for sets of dummy variables, with its p-value in parentheses below. Otherwise, standard errors are reported within the parentheses.

Table 9: Summary of Results

	1914 Importation	1917 Production	1923 US Production
Sample	All Dyes	Imported in 1914	Imported in 1914, Produced in at least one of 1917, 1918 & 1919, and Produced or Imported (but not both) in 1923
Source	Table 2, Column (7)	Table 4, Column (7)	Table 6, Column (10)
Discovery Year	Zero	Negative	Zero
Log (1914 Imports)	-----	Positive	Zero
Patents	Positive	Zero	Zero
# of Intermediates	Zero	Zero	Negative
Schultz Number	Negative	Negative	Zero
Dye Classes	XXX	XXX	XXX
Joint Technical	XXX	XXX	XXX
Colours	XXX	Zero	Zero
Materials	XXX	XXX	XXX
Joint Demand	XXX	Zero	XXX

Zero indicates that the corresponding t-test or F-test was insignificant. Positive (Negative) indicates that the corresponding t-test was significant and positive (negative). XXX indicates that the corresponding F-test was significant.