The Dye Famine and its Aftermath: Information Diffusion and Entry

David Genesove

Hebrew University of Jerusalem and C.E.P.R.

Initial Draft: May 18, 2005

This Draft: November 23, 2005

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, and Pierre Regibieau, Joachim Henkel and Julie Holland Mortimer for acting as discussants at the CEPR and NBER conferences in IO. Thanks to other participants at those conferences, as well as to seminar participants at the European University Institute, Bar-Ilan University, and the London Business School.

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. But if an informational advantage is the source of incumbency, then its greatest enemy is surely information diffusion. Knowledge leaks out of the firm, or develops independently elsewhere in the commercial realm, thus threatening the initial dominance.

Estimating the speed of information 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 information 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 information 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] See Vickers and Yarrow and for an early analysis of this issue. For a general treatment under myopia, see Athey and Schmutzler (2001). Increasing dominance is usually discussed in the context of strategic models, but it can also appear under perfect competition, as in Klepper, 1996.

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 of interest under the maintained assumption that the older the dye (i.e. the earlier the year of its discovery) the longer the time that information 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.

As 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 control for the missing variable through 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 information 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. Cantwell (19--) has shown how Germany remained dominant in the industry despite the destruction of much of its factories during World War II, which points to the crucial roles played by human and organizational capital in firms' success.

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 I use in this study. 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 thoroughly documented by the U.S. government. In aid of the deliberations over the 1916 tariff legislation, the Commerce Department collected and published figures on the quantity of dye imports (and, for a subset of them, their value), by the Schultz category (Norton, 1916).[2] (Although the identity of the exporting firms was recorded in that publication, firm level import quantities were only released in 1922.) 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 to 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), 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 (1995, 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.

The synthetic dye industry is usually dated to 1856, the year of Perkin's discovery in the UK 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 additional research and discoveries of a number of other dyes, and processes for the same, by other chemists, all of which led to massive patent litigation. The legal expenses and competitive effects of this litigation are generally believed to have been responsible for the industry's shift to the German states, where there was essentially no patent law. [3] This move was embodied in the homeward migration of a number of German chemists who had been residing in England. Among them was H. Caro, who in 1869 developed the first of the alizarin dyes (simultaneously with Perkin), the second major class of dyes to be developed, as well as A. Hoffman, Perkin's initially sceptical dissertation advisor. As in the case of the aniline dyes, this discovery led to the introduction of many other related dyes, and prices fell (Hohenberg). As a result (so the story goes), patent legislation was introduced in the newly formed Germany in 1877. 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. Germany's universities and their easy interrelationships with the industry appear to have been crucial to Germany's comparative advantage in this industry.

Figure 1 shows the distribution of dyes by the year of discovery[4] for the population of dyes that were 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. There is a sharp fall-off after 1910. Part of this surely reflects Schultz' ignorance in 1913 or 1914 of the very most recent discoveries, but not all, as *The Color Index*, (Rowe, 1924), shows a more moderate fall-off, with the number of discoveries in those years at only about a third or fourth of the immediately preceding years The distribution is relatively spread out, and has a single peak in the early 1890s.

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

German firms were extremely dominant in the industry. In 1913, on the verge of the war, they were producing between 80 to 90 percent of the dyes consumed in the world. The only other net exporter was Switzerland, which produced about ** of world production (Reader, 1970). In contrast, American firms produced only about two percent of world dyes, or about 13% of U.S. dye consumption. 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 dyes. Altogether, the 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" – most of the intermediates (about 240 out of a total 300, according to Haynes (Vol. 3, p. 212)) used in the production of the dyes in the United States were imported from Germany. 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. 123). 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. Only four and a half percent went to American firms.

In preparation for analysing the incidence of American production after the Dye Famine, it is useful to consider the determinants of pre-war imports. **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

[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." (Hearings, 1919.)

category, according to Schultz (1914), as an indicator of profitability. (See, for example, Putnam, 1997.) 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 0. The dyes in each class generally share a common tar-crude or even a basic intermediate. 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 dummy variables for the dye's colours. Column (6) adds dummy variables for the material (cotton, wool and silk) to which the dye may be applied.[6] Both sets of variables are highly significant. Finally, column (7) includes all the aforementioned variables. Each (set of) variable(s) remains significant (or 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 that embodied in the set of dyes available) and, through price, 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 neither demand nor costs.

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

Soon after the start of hostilities, Britain's navy blockaded Germany. This did not immediately cut off the supply of dyes to the United States, as there were domestic stocks to draw from, as well as stocks in China, Japan, Hong Kong, India, and the United Kingdom. Twice German dyes were delivered by submarine (thus evading the blockade). But the submarines carried a relatively small supply and it appeared to be essentially a publicity stunt (Steen). 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.

By April 1915, the stocks were drawn down. This marks the beginning of the Dye Famine, which was characterized by a low quantity of imports and so

[6] The source of the data for both sets of variables is *Colour Index*, 1924. Since this was published several years after the period in question, Schultz, 1914, which also presents this information, might be a preferable source.

consumption, high prices, and substitution to natural dyes. Figure 3 shows quarterly import quantities of Alizarin dyes. This is the only major dye category for which consistent import figures are available over this period. Here we see that imports plummeted between the first and second quarters of 1915. Although they subsequently recovered somewhat, import figures still remained an order of magnitude or two below its earlier level. **Figure 4**, 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.[7] **Figure 5**, from the same publication, shows that the price of natural dyes, an obvious substitute, rose dramatically as well.

The profit opportunities were obvious, and many U.S. firms entered the industry. By May 1916, if not earlier,[8] seventeen additional firms had entered the dyestuff industry (Norton, May 1916); by the time of the 1917 Census, some 81 firms were manufacturing synthetic dyes. The largest investor of them all was Du Pont, which had already been operating in the related explosives industry. But Du Pont was to run second in market share to National Aniline, the result of a merger of a number of vertically related firms, 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 strategy was to refer 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 (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 had no value. Hounshell and Smith (1988) relate, for example, how Du Pont managed to substantially improve the design for its plant to produce the important explosive intermediate diphenylamine based on a footnote from the *Journal fur pratische chemie*.

Of particular note was the inadequacy of the patent description. Apparently, the descriptions there were incomplete as well, with vital information on catalysts, as

[7] 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.
[8] It is not clear from the source whether the date is November 1915, the date of the first edition, or May 1916, the date of the revision.

well as optimal temperatures, pressures and timing missing.[9] 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 claim 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.

A second type of technology transfer was by way of human capital. In 1919 Du Pont tried to hire the leading German dye expert, Rene Bohn (inventor of the last new class of dyes, indathrene, in 1901). Although Bohn found the invitation to be appalling, and so declined it, some ten more junior chemists took up the offer the following year. One received $25,000/year, about $275,000/year in 2004 dollars and a tremendously large salary at the time. In contrast, Murmann, citing … Calco, in contrast, used a Yale chemistry professor (trained in Germany, of course). Human capital was important not only in the research end of the firms. The American firms hired the local marketing agents for the German firms, who would have had important information on demand, and buyers’ identities (Hounshell and Smith, p. 82; Steen, 1995, p. 25) ; 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).

A third method was to acquire the necessary information 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 previously existing dye firm, it had gained additional information on 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 information 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

[9] A dramatic enactment of this was provided in the 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, 19--). How long it would have taken him to get it right, if at all, can not be known.

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.[10]

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. 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 spent very close to two and a half million dollars on those labs, or about 3.6% of dye sales. 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).

Notwithstanding these various efforts, the Americans were clearly unable to fully replace the dyes that had been cut off from Germany. **Table 0** 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, 827 are ascribed discovery dates that are 1875 or later. Of these, 508 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 – 21, or six and a half 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

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

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.

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

Section IV: A Model for American Production

(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),

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

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)

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

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

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

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). I assume 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) $$\Pi_i = \gamma^{-1} sign(\gamma) S_i Z_i^{\gamma+1}$$

where we define $Z_i \equiv sign(\gamma)\frac{\gamma}{1+\gamma}(\alpha_i - c_i)$, 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:

$$\pi_i = s_i + (\gamma+1)z_i^U + e_i \; = \; (q_{1914,i} - \gamma z_i) + (\gamma+1)z_i^U + e_i$$

$$(8) \qquad = q_{1914,i} + z_i + (\gamma+1)(z_i^U - z_i) + e_i$$

where small letters indicate the log of the variable, the 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).

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 increases in either the numerator or the denominator will decrease (increase) z_i if γ is positive (negative), while its "coefficient", $(1-\varphi_i)$, is negative (due to the Germans' marginal cost advantage[11]), we should expect variations in this factor to impart a positive (negative) bias to the estimated coefficient on imports if γ is positive (negative)[12].

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 information diffusion effect. Thus without knowing the sign of γ we can not sign the bias. (It may be possible to determine γ's sign from the pass-through rate of differential tariff changes to price, since when γ is positive (negative), costs are passed through by less (more) than one hundred per cent.)

We proxy for the unobserved 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

---

[11] It is clear from all sources that the Germans faced a lower cost of production than the Americans. Taussig (1922) attributed it to the greater availability of skilled, technically educated workers in Germany. Schröter writes that specialized dyes cost two to three times as much when produced in a German subsidiary in the U.S. than in Germany, but that the factor fell for bulk dyestuffs (Schröter, p. 180). Murmann, citing …, states that on average the U.S. cost was about forty percent more than the German cost.

[12] 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.

procedures. That information is being coded up presently. In the meantime, we will use the number of intermediates used in each dye, from Shreve, 1922.

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 variation in the degree of concentration discussed below.

Note also that the American firms would have had limited information 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.[13] The 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 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 degree of concentration (which is unobserved) 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 $Var(h)/Var(q_{1914})$.

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

[13] Preliminary information was released a few months before, at the ? Exhibition (Steen, 1995).

variance of the log-price is therefore an upper bound for $Var(h)$:

$Var(p_{1914}) = Var(h) + Var(\ln[\alpha + \gamma c]) \geq Var(h)$. As noted earlier, the 1914 import price is not available for all dyes, but among those for which it is available (322 of the 508 imported), the variance of its log is .56, compared to 3.05 for the variance of log imports. (See Table 1.) Thus we can conclude that the estimated coefficient on log imports is biased downwards, and so $\hat{\beta}$ biased 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

$$U(q_1, q_2, ..., q_M) = G(\sum v(q_{g,j})) = G(v(\sum (\alpha_i - (\frac{q_i}{S_i})^{1/\gamma})))$$

[to be continued]

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

(5') $P_i = [\alpha_i + \gamma N c]/(1 + \gamma N)$

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

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

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

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 ($\gamma = 1$, $\kappa = 1/2$) 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 is to treat γ as an estimable parameter. That would be appropriate, however, 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 α and lower costs should 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).[14] 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.

---

[14] 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.

Since the Dye Famine cut off imports from Germany but not from Switzerland, which was the source of almost all non-German imports, we also include the total market share of Swiss firms.

Section V: American Production in 1917

Table 3 presents the linear probability regression of the incidence of US production in 1917. The sample is the 508 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.[15] 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 -.24 and is reported in the third to last row of the table. Its standard error, calculated directly from the corresponding nonlinear least squares model in which the coefficient on log imports is set to one, and the scale of the uniform distribution, divided by ρ, is estimated, is reported on the row beneath. The estimated ratio has the interpretation that a dye that is one year "younger" will cost 1-exp(.24) = 27 percent more to develop. This implies, in turn, that it was more than 9000 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

[15] 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),

noting at this point that these numbers, although large, are not completely unreasonable. It is not inconceivable that information 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 information 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 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 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.[16] "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 patent out is added. Both variables are insignificant, both singly and jointly, indicating that, not withstanding the claims of the US firms at the time, claims that 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. 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.

The remaining columns control for market structure in 1914. Column (4) adds the inverse of the number of German firms exporting to the U.S. in that year.[17] We have noted that the endogeneity of the number of firms is likely to bias its estimated coefficient, and indeed it is highly negative, rather than positive, as would be expected

[16] 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 certain assets, but not patents. Seizing the latter was permitted only by an amendment to the act in November 1918, right before the armistice.

[17] By 1914, the domestic sales operations of the six (eight?) largest German companies were organized in two cartels. There is some question .thesis argues that the firms competed in the United States.

were the variable exogenous. Column (5) adds the HHI, which generalizes the inverse number of firms to the case of non-equal market shares. Now, its is coefficient is significantly negative, whereas the inverse number of firms is insignificant. Columns (6) and (7) add the 1914 market share of Swiss firms; its coefficient is insignificant in both cases. (Note that both the HHI and the Swiss market share rely on information published only in 1922). The estimated ratio $\hat{\beta}$ is robust to the inclusion of all three of these variables.

Table 4 adjusts the log quantity of imports to account for the presence of more than one firm, using the Cournot model, as described above. We present 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 Table 5 below. 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

An alternative explanation for the negative coefficient on discovery year is intrinsic differences in the difficulty of development, which we will term complexity. **Table 5** considers the inclusion of 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 had to determine how to produce more intermediates, and (b) there were more chemical steps involved. Like all proxies for complexity, however, this is likely to effect not only development but also production costs, although increases in either would make development and production less likely. As column (1) (which should be compared to column (2) of the previous table) shows, inclusion of this variable has no effect on either the coefficient on discovery or that on log imports. However, the estimated coefficient on the new variable itself is insignificant, throwing doubt on the adequacy of the proxy.

Column (2) considers the set of 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. Controlling for them has a clear effect on the coefficients of interest, decreasing the magnitude of both, but of the discovery year 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 21 percent more to develop, and that the 1913 cohort will cost almost 1400 times than the 1875 cohort – a factor much more in line with the rates from Beer and Gambaredella given above.

Column (3) adds the Schultz number itself (again, normalized to lie between zero and one). It is highly significant, and decreases $\hat{\beta}$ in magnitude to -.17. One can only guess why the Schultz number is relevant to the analysis. As **Figure 6** 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 information 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.

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.

With demand attributes we face no such problem of interpretation. Columns (4) through (6) adds the set of colour dummy variables and the set of materials, first singly and then together. The F-tests show the set of colours to be only marginally significant, but the materials highly significant (overwhelmingly due to the negative effect of cotton). But when we add the "technical" attributes back in, these two sets of 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 (8) adds the inverse number of firms. It again has a negative coefficient. This time it affects $\hat{\beta}$, reducing it two points in magnitude to -.15.

Finally, Table 6 adds the log of the import price.[18] That cuts down the number of observations by almost 40 percent, as this variable is not always reported in Schultz (1914). No matter what other regressors are included , price is highly insignificant and has almost no effect on $\hat{\beta}$. The general pattern of significance of the demand and technical attributes remains and, indeed, is even starker now, with the set of materials variables always insignificant.

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

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 information 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 information 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.[19] However the bias will be small if most of the

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

[19] 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.

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 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. It is possible that the third set might work where these have not, but I have not succeeded yet in obtaining the requisite data to construct such instruments.

As an alternative procedure, I restrict the sample further by including only those dyes that (a) were produced by American firms in 1917-1919, 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. However, e (to recall, the horizontal demand shock in 1917, relative to 1914) and η are potentially correlated; it is thus the presence of η that generates the selection bias. Note the implicit assumption, then, that variations in demand affect both domestic and foreign firms equally. This will be true for a horizontal demand shifter, as e is defined to be, and is consistent with differences in monopoly profits arising from differences in cost.

The probability, conditional on factors X and the common shock 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. 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 7 shows the results.[20] (For the sake of comparison with the earlier tables, and without apology, it employs a linear probability model, instead of a logit.) The major finding here 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

[20] 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 and with the variables thus redefined, we obtain similar results to Table 6.

effect of information 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. 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 $\sqrt{508/95}$ - the square root of the ratio of the number of observations in the imported dyes and this sample), 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.

Considering columns (3), (4) and (6), where technical variables are added, singly or jointly, but without demand variables, we see some indication that more complex dyes are more likely to be imported than produced by the American firms: the estimated coefficients on the number of intermediates and the Schultz number are either negative and significant or insignificant. However, when in Column (10) demand variables are added, those two variables take on opposite signs. However, Column (10) may be somewhat suspect given the large number of coefficients (33) relative to the number of observations (94). The set of class dummies is jointly significant. As for the demand variables, which appear in columns (7) through (10), the set of colours are insignificant while the set of materials is (generally) significant.[21]

Section VII: Conclusion

[Also preliminary]

The central finding in this paper is that the greater the number of years since a dye's discovery, the more likely that American producers were to produce the dye. When compared to the sensitivity of the incidence of production to the log of imports, this implies a rate of information diffusion, where information is measured in terms of

[21] This is consistent with the above model and previous results if colours affect horizontal demand shifters (whether via s or e), whereas materials affect vertical demand shifters (α). Because US and foreign firms have different costs, variations in α will affect log monopoly foreign and domestic profits differentially. But that may be taking the specifics of the model too seriously.

the additional cost necessary for development. The actual rates measured range from 14 percent to 26 percent a year, depending on the specification.

The pattern of signs and significance of these and other variables, which is summarized in Table 8, buttress this interpretation. The discovery year has no effect on importation in 1914, while demand and cost proxies have strong effects, indicating that the discovery proxies for neither. In contrast, the discovery year has a negative effect on US production in 1917, indicating that it increases the fixed cost of development.

Among those dyes produced by the Americans before the re-entry of the Germans, and then either produced or imported in 1923, discovery year again has no effect on U.S. production – suggesting that is has no effect on the relative cost of production. Together, these findings strongly suggest that the discovery year is an indicator of information diffusion, which is necessary for a dye's development, but irrelevant for production efficiency.

[to be continued]

REFERENCES

Note: Not all bibliographic items are referred to in the text yet.

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.

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.

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

Cantwell, “The Evolution of European Industrial …”

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

Cohen, W.M. and D.A. Levinthal (1989): "Innovation and learning: The two faces of R&D," *Economic Journal*, 94, 569-596.

Cohen, Nelson, and .

Cohen, Welsh and Walsph, N.B.E. R. Working Paper, February 2000.

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.

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 and Mullin, ‘Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914”, *RAND Journal of Economics*. Summer 29(2): 355-, 1998.

Griliches and Yatchew, “Specification Error in Probit Models”, *Review of Economics and Statistics*, 134-139, 198x.

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.

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.

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

Mansfield, M. Schwartz and S. Wagner, "Imitation Costs and Patents: An Empirical Study," *Economics Journal* 91 (1981) 907-918.

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.

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

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

Reader, , *Imperial Chemical Industries:* , 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.

Schultz, Gustav. *Farbstofftabellen*, (5th edition) 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*, August 1995, Dissertation, University of Delaware.

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

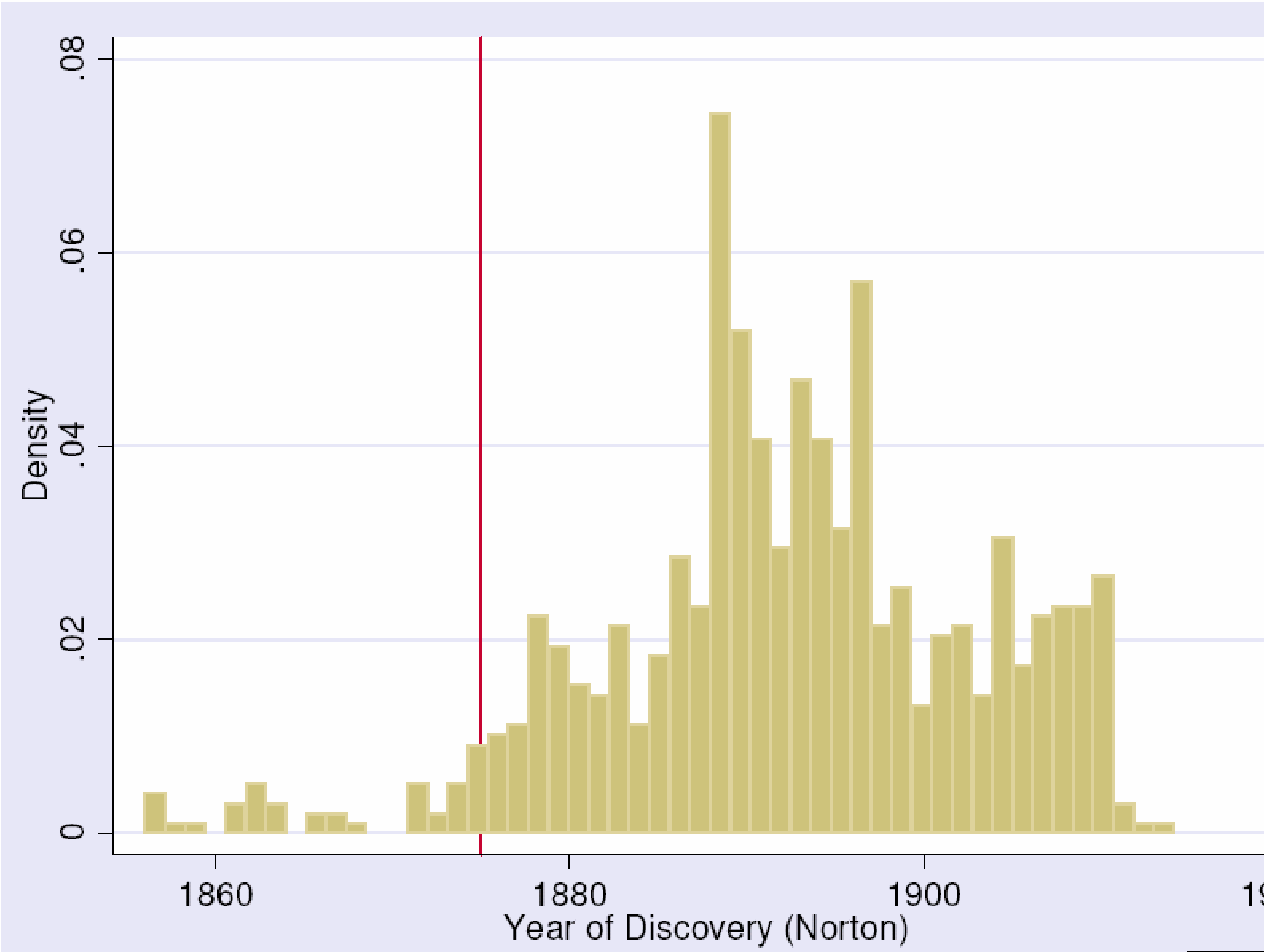

FIGURE 1: The distribution of dyes by year of discovery.

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

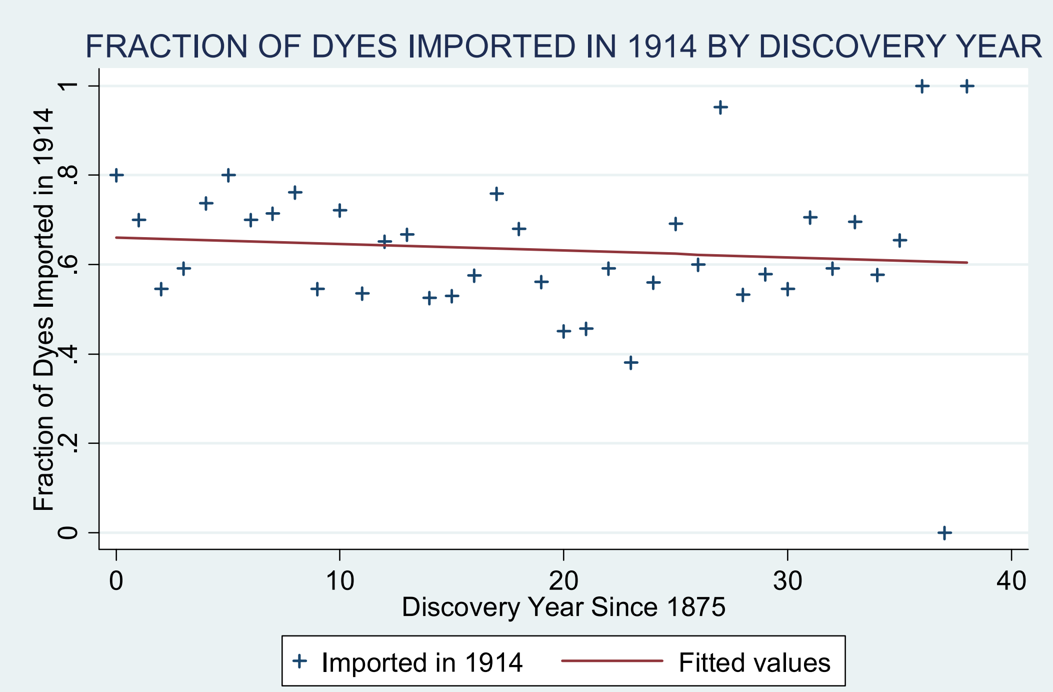

FIGURE 2a

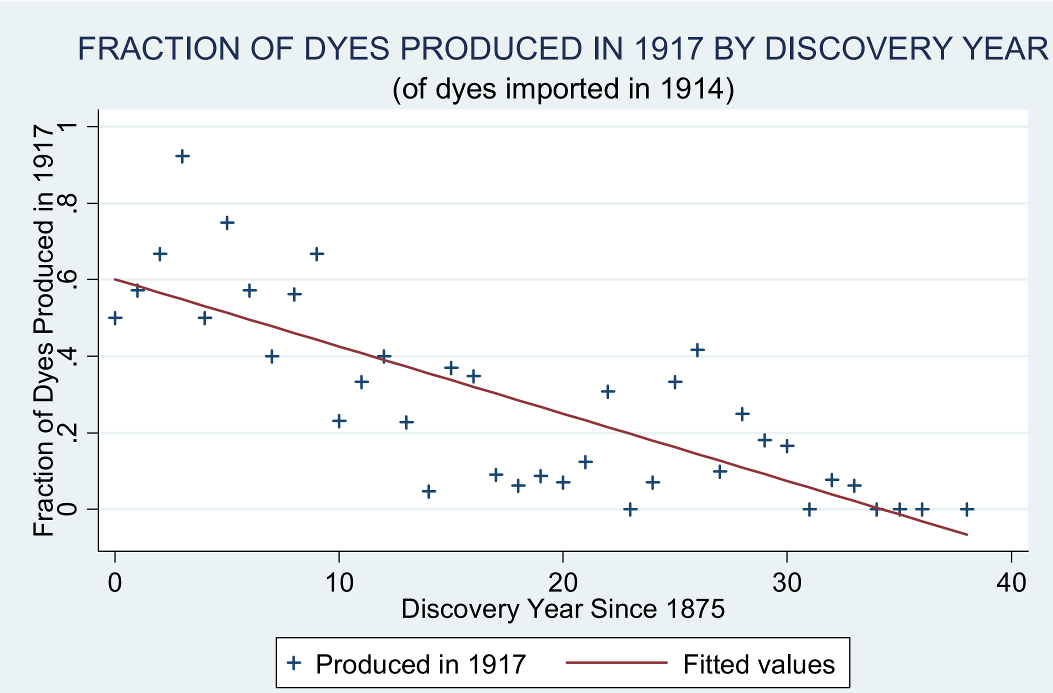

FIGURE 2b

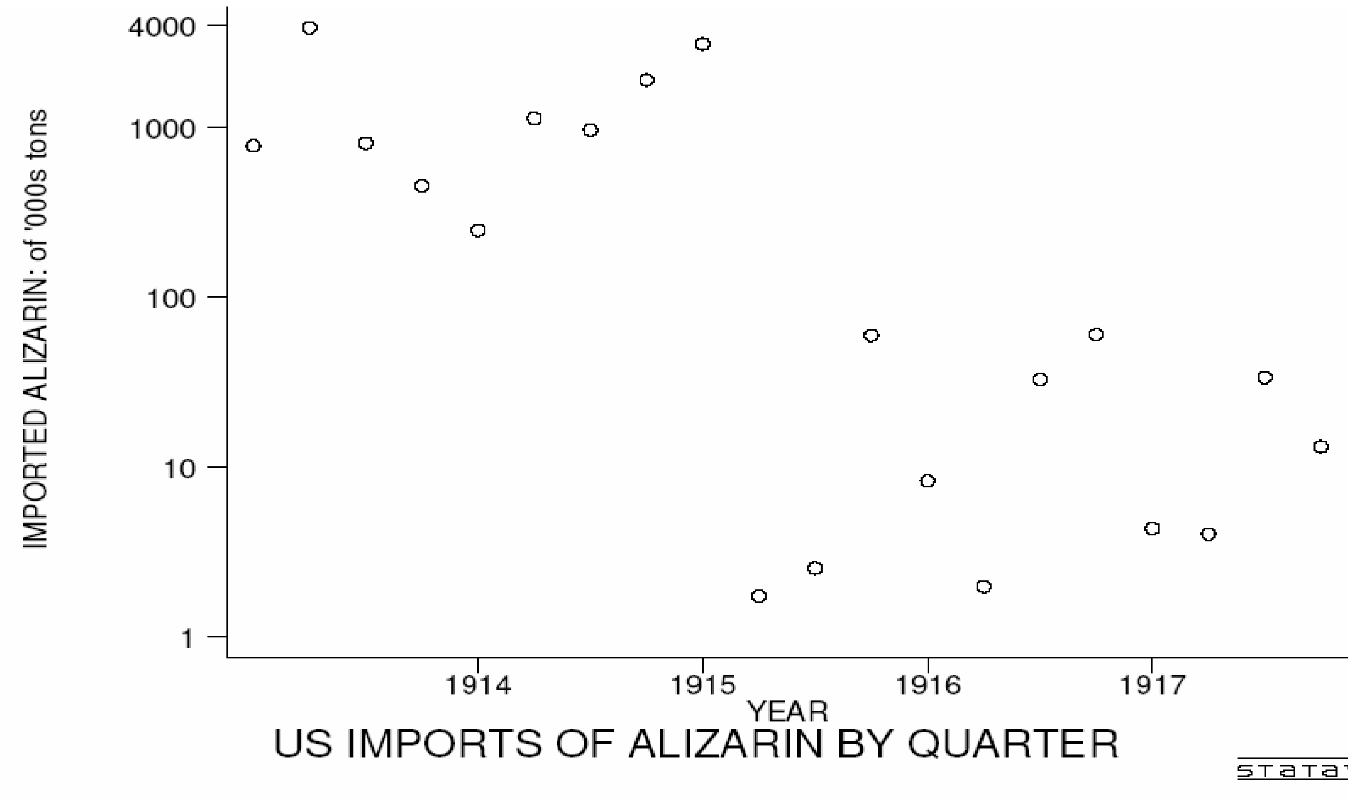

FIGURE 3: U.S. Imports of Alizarin by Quarter

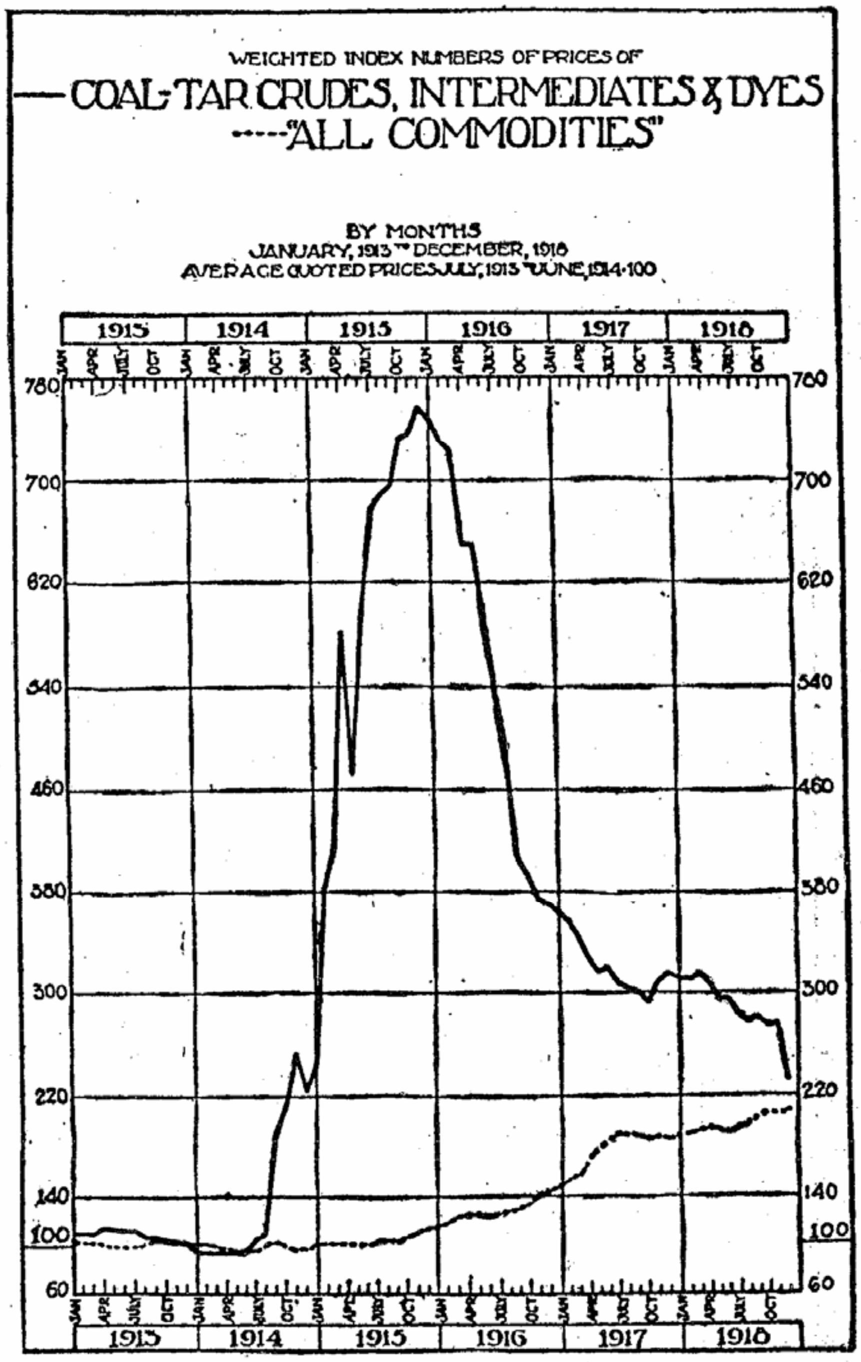
WEIGHTED INDEX NUMBERS OF PRICES OF
— COAL-TAR CRUDES, INTERMEDIATES & DYES
----"ALL COMMODITIES"
BY MONTHS
JANUARY, 1913 TO DECEMBER, 1918
AVERAGE QUOTED PRICES JULY, 1913 TO JUNE, 1914 = 100
1913
1914
1915
1916
1917
1918
JAN
APR
JULY
OCT
780
700
620
540
460
380
300
220
140
100
60

FIGURE 4

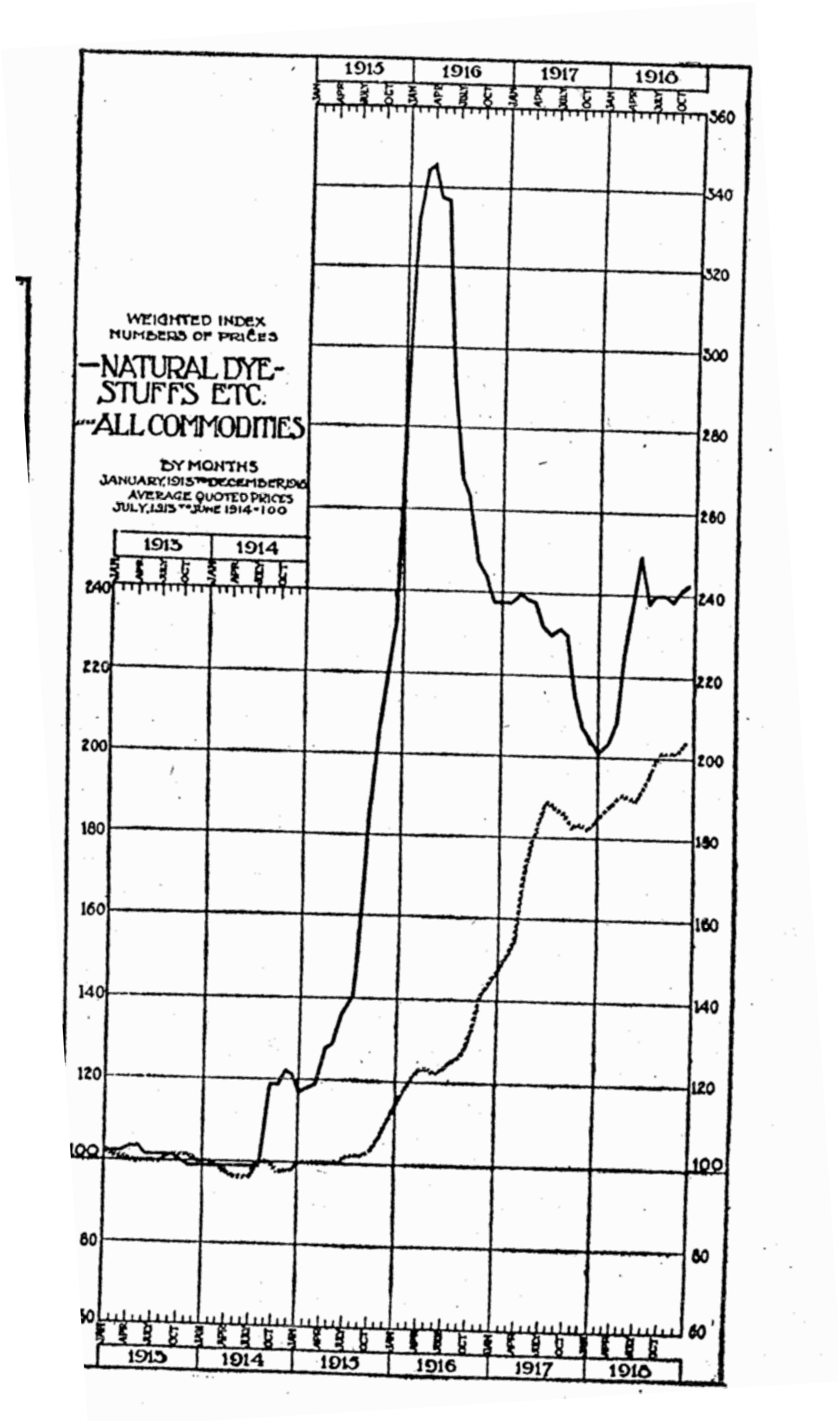
WEIGHTED INDEX NUMBERS OF PRICES
—NATURAL DYE-STUFFS ETC.
···ALL COMMODITIES
BY MONTHS
JANUARY, 1913 TO DECEMBER 1918
AVERAGE QUOTED PRICES
JULY, 1913 TO JUNE 1914 = 100

FIGURE 5

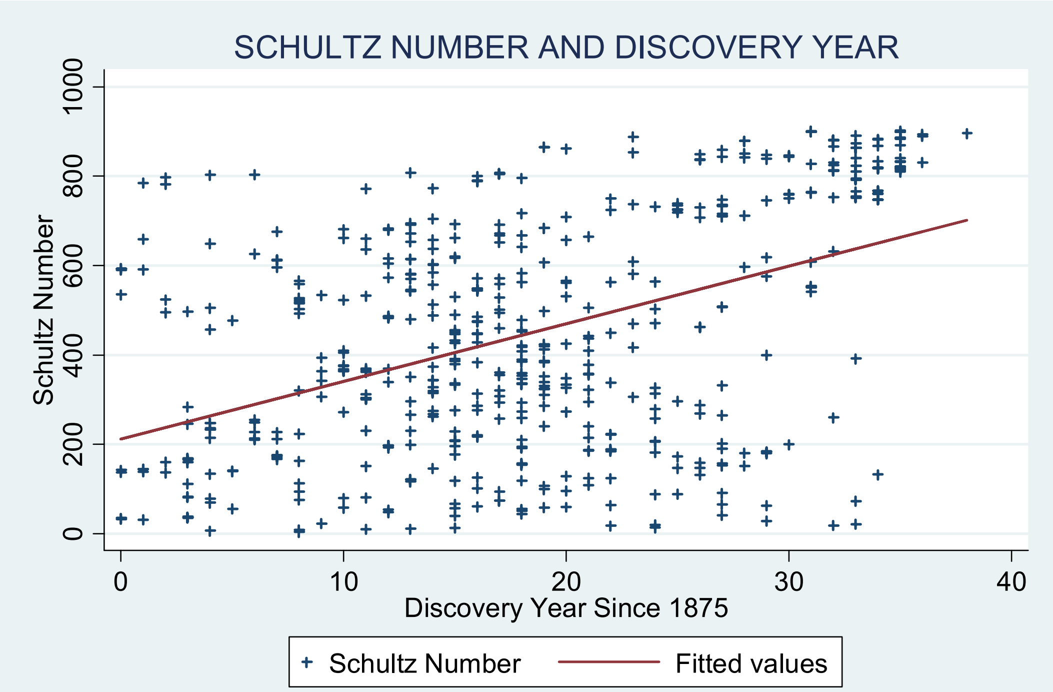
SCHULTZ NUMBER AND DISCOVERY YEAR
Schultz Number
0
200
400
600
800
1000
0
10
20
30
40
Discovery Year Since 1875
Schultz Number
Fitted values

FIGURE 6

Table 0: Importation in 1914 and American Production in 1917

	Not Produced in 1917	Produced in 1917	All
Not Imported in 1914	298	21	319
Imported in 1914	385	123	508
All	683	144	827

Table 1: Summary Statistics

Sample:	Imported in 1914 (N=508)		Imported in 1914, and Value Imported Reported (N=322)	
	Mean	Standard Deviation	Mean	Standard Deviation
Discovery Year – 1875	18.37	9.34	16.87	9.38
Quantity (pounds)	40,114	93,88	57,239	109,619
Ln Quantity	8.09	2.08	9.73	1.75
Price (dollars)			0.31	0.33
Ln Price			-1.46	0.75

Table 1: Summary Statistics

	All post 1875 Dyes (N=827)		Imported in 1914 (N=500)		Imported in 1914 & Price reported (N=312)		Produced by US in 1917-20 & prod. / imp. 1923 (N=95)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Discovery Year - 1875	18.52	9.08	18.37	9.34	16.87	9.38		
Ln (Import Quant)			8.89	2.08	9.73	1.74		
Import Quantity (tonnes)			40.1	93. 9	57.2	109.6		
Ln Import Price					-1.46	0.75		
Price (dollars)					0.33	0.31		
Schultz Number	0.49	0.28	0.48	0.29	0.46	0.28		
No. of Intermediates	2.44	1.01	2.45	1.07	2.48	1.03		
	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: American Imports in 1914

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Discovery Year – 1875	-.001	-.001	-.002	-.001	-.002	-.0002	-.002
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
# Countries with Patents		.020					.027
		(.012)					(.011)
16 Dye Classes			2.43				2.17
(p-value)			(.002)				(.006)
Schultz Number				-.03			-.40
(Normalized on [0,1])				(.07)			(.19)
16 Colours					3.17		3.02
(p-value)					(.0000)		(.0001)
3 Materials (Wool, Cotton,						4.59	3.23
Silk) (p-value)						(.003)	(.022)
Constant	.63	.59		.64			
	(.04)	(.05)		(.04)			
Classes & Schultz Number							2.07
(p-value)							(.008)
Colours & Materials							3.16
(p-value)							(.0000)
R-squared	.0005	.004	.044	.001	.060	.018	.121
Number of Observations.	817	817	817	817	817	812	812

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: American Production in 1917

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Discovery Year – 1875	-.017	-.015	-.015	-.010	-.010	-.015	-.010
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Ln(Import Quantity 1914)		.063	.062	.041	.042	.063	.040
		(.008)	(.008)	(.010)	(.010)	(.008)	(.010)
US Patent in Force			-.004				
			(.052)				
# Countries with Patents			.009				
			(.017)				
1 / Number of Exporters				-.27	-.004		-.01
				(.06)	(.11)		(.10)
HHI					-.42		-.41
					(.13)		(.13)
Swiss Market Share						-.04	-.04
						(.05)	(.05)
Constant	.55	-.05	-.06	.004	.40	-.04	.41
	(.04)	(.09)	(.09)	(.08)	(.12)	(.09)	(.12)
R-squared	.13	.23	.23	.25	.27	.23	.27
Ratio of Coefficients ($\hat{\beta}$)	----	-.23	-.24	-.25	-.24	-.23	-.24
	(---)	(.04)	(.05)	(.07)	(.07)	(.04)	(.07)
Number of Observations	500	500	500	499	499	500	499

The sample in Columns (1) through Columns (5) is all dyes that were imported in 1914 for which a discovery year of 1875 or later is available. Column (6) restricts the sample to those dyes for which a German firm was one of the exporters. Column (7) restricts the sample to those dyes for which a Swiss firm was one of the exporters. Standard errors are reported within the parentheses. (The discrepancy between 500 and 508 is explained by some duplicate observations, due to merging. This will be corrected in a future draft.) Ratio of Coefficients ($\hat{\beta}$) is the ratio of the coefficient on year of discovery to the ratio of the coefficient on Ln(Import Quantity). The standard error is calculated by the delta method, as implemented in Stata's nlcom command.

Table 4: Adjusting Quantity Imports for the (Equivalent) Number of Firms

		(1)	(2)	(3)	(4)
Control Variables	1/N or HHI	$\gamma = 0$	$\gamma = 1$	$\gamma = \infty$	$\gamma = -1.1$
None	1/N	-.24	-.25	-.26	-.36
		(.04)	(.04)	(.04)	(.08)
Yes	1/N	-.17	-.17	-.18	-.27
		(.04)	(.05)	(.05)	(.07)
None	HHI	as above	-.25	-.26	-.36
			(.05)	(.05)	(.08)
Yes	HHI		-.17	-.18	-.27
			(.04)	(.05)	(.07)

Table 5: American Production in 1917

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Discovery Year – 1875	-.015	-.012	-.011	-.014	-.014	-.013	-.012
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Ln(Import Quantity)	.064	.065	.066	.067	.062	.066	.069
	(.008)	(.008)	(.008)	(.009)	(.008)	(.009)	(.009)
Number of Intermediates	-.010		-.016				-.010
	(.017)		(.020)				(.020)
16 Dye Classes		2.03	1.83				1.73
(p-value)		(.012)	(.03)				(.04)
Schultz Number			-.77				-.55
			(.17)				(.20)
14 Colours				1.52		1.44	.95
(p-values)				(.09)		(.13)	(.50)
3 Materials (Wool, Cotton,					5.86	5.52	1.93
Silk) (p-value)					(.0006)	(.001)	(.12)
#Intermed. & Classes &			*3.37*				*2.21*
Schultz # (p-value)			*(.0000)*				*(.004)*
Colours & Materials						*2.19*	*1.08*
(p-value)						*(.003)*	*(.37)*
R-squared	.22	.27	.31	.26	.25	.29	.34
Ratio of Coefficients ($\hat{\beta}$)	-.23	-.19	-.17	-.21	-.22	-.20	-.17
	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)	(.04)
Number of Observations	493	500	493	500	497	497	490

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 6: American Production in 1917 (with log-Price added)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Discovery Year – 1875	-.017	-.013	-.013	-.017	-.016	-.016	-.015	-.009
	(.002)	(.003)	(.002)	(.003)	(.003)	(.003)	(.003)	(.004)
Ln(Import Quantity)	.071	.080	.084	.081	.070	.080	.090	.065
	(.014)	(.014)	(.014)	(.015)	(.014)	(.015)	(.016)	(.017)
Ln(Import Price)	-.049	-.002	.003	-.030	-.045	-.029	.003	-.002
	(.033)	(.038)	(.038)	(.035)	(.034)	(.036)	(.040)	(.039)
Number of Intermediates	-.0004		-.009				-.004	-.013
	(.024)		(.027)				(.028)	(.028)
16 Dye Classes		1.71	1.72				1.77	1.79
(p-value)		(.05)	(.05)				(.04)	(.04)
Schultz Number			-.81				-.73	-.53
			(.23)				(.29)	(.20)
14 Colours				1.07		1.04	.81	.63
(p-values)				(.04)		(.41)	(.68)	(.84)
3 Materials (Wool, Cotton,					1.62	1.46	.92	1.35
Silk) (p-value)					(.19)	(.23)	(.43)	(.26)
#Intermed. & Classes &			*2.35*				*1.99*	*1.88*
Schultz # (p-value)			*(.002)*				*(.01)*	*(.02)*
Colours & Materials						*1.14*	*.85*	*1.01*
(p-value)						*(.31)*	*(.64)*	*(.45)*
1/Number of Exporters								-.34
								(.09)
R-squared	.24	.30	.33	.28	.25	.29	.34	.36
Ratio of Coefficients ($\hat{\beta}$)	-.24	-.20	-.16	-.21	-.23	-.20	-.17	-.14
	(.06)	(.05)	(.05)	(.06)	(.06)	(.05)	(.05)	(.06)
Number of Observations	312	316	312	316	314	314	310	309

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 7: US Production in 1923, among dyes that were (a) produced in 1917-19 & (b) produced or imported in 1923

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Discovery Year – 1875	-.0012	-.0013	-.0003	-.0008	-.0018	-.0004	-.0000	.001	.003
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)	(.005)	(.004)	(.005)
Ln(Import Quantity)		.019	.020	.018	.022	.023	.011	.016	.007
		(.016)	(.016)	(.015)	(.016)	(.016)	(.018)	(.016)	(.012)
# of Intermediates			-.04			-.11			
			(.05)			(.05)			
Schultz Number				-.44		.30			
				(.17)		(.29)			
12 Dye Classes					2.63	2.37			
(p-values)					(.005)	(.01)			
14 Colours							1.63		1.25
(p-values)							(.10)		(.26)
3 Materials (Wool, Cotton, Silk) (p-value)								3.99	2.38
								(.01)	(.08)
R-squared	.002	.017	.03	.04	.29	.30	.21	.13	.28
Number of Observations	95	95	95	95	95	95	95	94	94

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 8: 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	
Colours	XXX	Zero	Zero
Materials	XXX	XXX	XXX
Joint Demand		Zero	

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.