Can Stimulating Demand Drive Costs Down? World War II as a Natural Experiment

François Lafond, Diana Greenwald, and J. Doyne Farmer

U.S. military production during World War II increased at an impressive rate and led to large declines in unit costs. However, the literature has focused on elucidating detailed mechanisms behind this relationship, using small datasets on specific products. Here we take a step back and, looking at an unprecedently large collection of data, we show that both exogenous technological progress and endogenous effects from increasing production experience were important, in roughly similar proportions. The demand for military products was largely exogenous, and the correlation between production, cumulative production, and time was weak, limiting issues of reverse causality and multicollinearity.

World War II has been the subject of a vast historical literature that ranges from first-person narratives of a handful of infantry soldiers on the frontlines to sprawling histories that chart the years-long movements of armies and interactions of now-famous heads of state. Under this large umbrella of the history of WWII, a more limited economic history has been written. Some works are overviews of how wartime economies functioned during the conflict (see, e.g., Tooze 2006; Herman 2012), while others are very narrow histories of the production of one product for the war effort, such as Liberty Ships (Thompson 2001, 2007). The economic history of WWII has also been appended to the macroeconomic histories

The Journal of Economic History, Vol. 82, No. 3 (September 2022) © The Author(s), 2022. Published by Cambridge University Press on behalf of the Economic History Association. doi: 10.1017/S0022050722000249

François Lafond is Senior Research Associate, Institute for New Economic Thinking at the Oxford Martin School and Mathematical Institute, University of Oxford, INET, Manor Road Building, Manor Road, Oxford, OX1 3UQ. E-mail: francois.lafond@inet.ox.ac.uk (corresponding author). Diana Greenwald is Assistant Curator, Isabella Stewart Gardner Museum, Boston, 25 Evans Way, Boston, MA. E-mail: dgreenwald@isgm.org. J. Doyne Farmer is Professor, Institute for New Economic Thinking at the Oxford Martin School and Mathematical Institute, University of Oxford, and Santa Fe Institute, INET, Manor Road Building, Manor Road, Oxford, OX1 3UQ. E-mail: doyne.farmer@inet.ox.ac.uk.

This work was supported by the EU grant FP7-ICT-2013-611272 (GROWTHCOM), Partners for a New Economy, Baillie Gifford, the EU grant 730427 (COP-21 RIPPLES), the Rebuilding Macroeconomics project (which is supported by the Economic and Social Research Council), and the Institute for New Economic Thinking at the Oxford Martin School. We would like to thank Kieran Marray for excellent research assistance and Nick Chater, Giovanni Dosi, Cameron Hepburn, David Hendry, Christopher Magee, John Muellbauer, Bill Nordhaus, Matteo Richiardi, our INET colleagues as well as audiences at the CCS 2016, EHS 2018, EEA-ESEM 2019, Singapore NUS, Oxford, MIT, LSE, OECD, CONCORDI 2019, and Waterloo for comments. The usual disclaimer applies.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

of fiscal stimuli applied by governments during the Great Depression; the war has been understood as a major source of government spending and therefore economic stimulus (Fishback 2020). In this paper, we seek to bridge the gap between these macrohistories—the large-scale narratives about the effects of the war on national economies—and the more tightly focused microeconomic histories of individual projects. We seek to do this by engaging with the study of learning-by-doing that frames the histories of specific wartime products, but have assembled and analyzed a dataset about far more products than any previous study.

An old and well-established literature demonstrates empirically that for many products, unit cost drops as cumulative production increases (Thompson 2012). This relationship has been called several different names, including the *learning curve*, the *experience curve,* and *Wright's law*. Wright's law is named for Theodore Wright, who conjectured that the cost of manufacture of a given model of an airplane by a given plant decreases as a power function of the cumulative number of planes produced (Wright 1936). Apart from Wright's observations, which were made in the 1930s, many experience curve studies focus on a distinct historical moment: WWII. (In fact, Wright was a policy advisor during the war effort.) This paper returns to that decisive moment, but with two novel contributions: new and far broader data about industrial production during the war and a more critical eye toward experience curves and their value in predicting declines in cost.

This observed relationship between costs and production has often been understood as causal and used to set policy. Large-scale public support for a particular industry is often justified by the argument that artificially increasing production will "push it down its learning curve" until it becomes competitive in a global market.[1] A critical assessment is now even more essential, as experience curves play a particularly important role in integrated assessment models for climate change mitigation, where they make it possible to model endogenous technological progress (Nordhaus 2014; Witajewski-Baltvilks, Verdolini, and Tavoni 2015).

There are three major related issues with using experience curves as a justification for policy (Thompson 2012; Nordhaus 2014). First, there is the possibility of reverse causality. Experience curves, at least when used to understand the effects of policy, postulate a causal relation from cumulative production to costs. However, as long as demand is elastic and costs are going down, lower prices can induce higher demand and

[1] The existence of learning curves has been a key argument for the protection of "infant" industries, typically through import protection using tariffs. The evidence that such policies are useful remains mixed (Lane 2020).

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

thus a rise in production and experience. A simple regression of costs on experience, therefore, does not only identify the causal effect of experience on cost but also picks up the effect of demand elasticity.

A second issue is that experience curves are plagued by an omitted variables bias. Productivity improvements take place for a variety of reasons. One of these is learning-by-doing, which refers to the improvements in the manufacturing process that are stimulated by production experience. We may think of experience as a proxy for all relevant product-specific cumulated efforts, and—absent the reverse causality issue mentioned previously—identify their causal effect on cost decrease. However, productivity improvements can also occur independently of product-specific or industry-specific experience or innovation, such as knowledge spillovers from other industries or the general improvement in a country's economic institutions.

Attempts to address this second problem of omitted variable bias lead to a third problem. The simplest option to control for technological progress that is *not* related to experience is to include a time trend, which is intended to capture all productivity improvements that are due to "exogenous" factors. In practice, however, adding a time trend to a learning curve regression introduces a serious estimation problem. In most empirical examples production grows exponentially, so that experience also grows exponentially. Because experience is the sum of past production, fluctuations are dampened. This means that the log of experience is similar to a deterministic linear time trend, and multicollinearity makes it impossible to distinguish the relative effects of experience from those of the time trend. While multicollinearity is well understood and does not by itself bias the coefficient of interest, standard errors become very large and coefficients are unstable under changes in specification, yielding ambiguous results. Focusing on the usefulness of experience curves for government support of climate mitigation technologies, Nordhaus (2014), in particular, showed that exogenous technological progress and reverse causality can lead to substantially biased and wildly uncertain estimates.

Military production during WWII provides a unique natural experiment (Thompson 2012), which helps us solve all three of these issues. Demand was largely exogenous because it was driven by battlefield needs. Because battlefield requirements quickly ramped up, then eventually plateaued and decreased, production, experience, and time are *not* strongly correlated. This allows us to estimate their relative effects more precisely.

Despite these advantages, many of the previous studies on WWII data (e.g., Thompson 2001) do not systematically control for exogenous

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

technological progress, and when they do, the results vary widely, with estimates of the effects of exogenous time trends being positive, negative, or insignificant. Instead, earlier studies have examined very specific datasets that often focus on one narrow area of WWII military production (Alchian 1963; Rapping 1965), and explained their value in the context of causal inference (Thompson 2001). As a result, much of what we know about military production in the United States during WWII is siloed in studies of specific product categories. In short, we have a great depth of information about certain products, but the economic history literature provides limited amounts of breadth.

Therefore, in this paper, we take a step back. Rather than trying to elucidate deeper causal mechanisms, we make one simple distinction—endogenous and exogenous effects—and gather as much data as possible to produce a consistent estimate. We find that endogenous and exogenous effects each contribute about half of the cost reductions, providing a direct, simple, and clear answer to Nordhaus's criticism of learning curves. In doing so, we show that historical analyses are useful for guiding policies for large-scale investments in critical technologies, like climate mitigation—an area in which experience curves are often cited with reference to providing government subsidies.

During WWII, the U.S. Armed Forces transformed from the world's 18th largest military to by far the world's largest (Herman 2012). The United States produced war equipment not only for themselves but for all their allies as well, manufacturing roughly two-thirds of the total equipment produced by both sides combined. This is an enormous exogenous demand shock. While the relative allocations to specific types of war equipment may have been somewhat cost-sensitive, the overall build up was clearly driven more by needs on the battlefront than by the price of weapons (Hall 2009; Ramey 2011). Moreover, capacity was often stretched to a maximum, further limiting endogeneity concerns that may have arisen if fast-learning producers were able to obtain more future contracts.

The massive build-up in military production was then followed by a dramatic slowdown. This makes the time series for production far from exponential and alleviates the problem of multicollinearity. Figure 1 illustrates a good example. Production of Ford's armored car, model M-20 GBK, began in May 1943 and increased to a peak of 400 units per month by September 1943; it dropped dramatically by spring 1944, rising later but never to the previous level. In contrast, costs came down more or less continuously, and a plot of the logarithm of cost against the logarithm of experience yields an approximately straight line. The key point

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

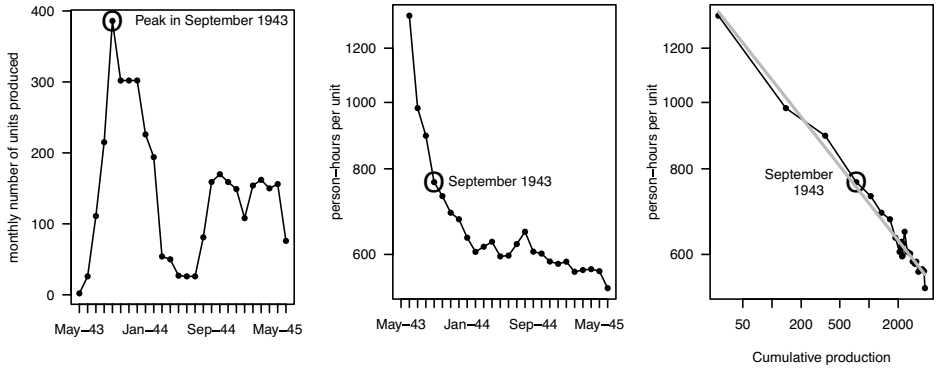


FIGURE 1
PRODUCTION (NUMBER OF UNITS PER MONTHS), PERSON-HOURS PER UNIT AND EXPERIENCE CURVE, USING CUMULATIVE PRODUCTION (WITH LINEAR FIT) FOR FORD'S ARMORED CAR M-20 GBK

Source: Ford's archives.

here is that costs continued to drop even as production *decreased*, lending evidence to the importance of *cumulative* production (experience) rather than production itself as a key variable (Asher 1956, p. 87).

To take advantage of this unique demand shock, we have gathered novel and unprecedentedly large amounts of data on U.S. military production during the war: time series on labor productivity, a cross-sectional dataset of unit cost declines for more than 500 products, and aggregate data on indices of contract prices broken down into 10 different categories.

These three datasets differ in terms of the main variable available (labor productivity vs. unit cost), number of observations (complete time series or not), level of aggregation (from plant to War Department level), and type of product. However, in combination, they provide an unprecedentedly comprehensive picture of industrial military production during WWII. While earlier studies have focused exclusively on complex "big ticket" items like ships and planes, our data captures the full gamut of war production: from mature technologies like rifles for the infantry to brand new technologies like early radar. To date, documentation of this diversity of production has been scattered across primary and secondary sources. Bringing it together represents a major contribution to the economic historical literature about WWII. In fact, we find that the rates of unit cost reductions were in general much smaller than what one could have thought based on the well-known examples of labor productivity growth in airplanes. This is likely to be because nonlabor costs, such as capital, became more prevalent, but also because airplanes were an exceptional category.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

To further bolster the value of this data, we have applied new corrections for estimating cumulative experience prior to the beginning of production observed in our data. Generally, earlier studies have simply applied a correction based on the assumption that production grows exponentially both within the observed sample and the period before observation starts. Because exponential growth would be a poor assumption in our case and because our datasets cover highly heterogenous products, we developed an alternative method of estimating prior experience, using historical statistics that are available at a coarse level to correct for prior experience at a finer-grained level.

Our results make it clear that costs depend on both experience and an exogenous time trend, in roughly equal amounts. Thus, experience *does* cause cost improvements, lending support to arguments that state-sponsored demand can be effective to accelerate the development of critical technologies. Nonetheless, even in a historical moment of enormous government investment, some ultimate causes of declining costs are better proxied by the passage of time than by the growth of experience.

Unfortunately, our data does not allow us to evaluate deeper channels of causality. For instance, higher productivity that follows a rise in experience can be due to direct learning-by-doing or to any productivity-enhancing factor that tends to scale with experience, such as R&D. We refrain from commenting on these specific channels, and simply conclude that endogenous effects (as captured by experience at producing a specific product) and exogenous effects (as captured by a time trend or time dummies) explain about as much of the average rate of cost decrease. As a result, a large demand shock is likely to accelerate declines in costs of production.

An important contribution to historical studies of WWII is that the relationship between experience and cost is very heterogeneous: while for some products, costs decline sharply if experience increases, other products are less sensitive to experience effects. As a result, we conclude that a policy stimulus is only likely to be effective for a particular product when there is a favorable historical relationship.

LITERATURE REVIEW

We provide a review that focuses on the points of overlap between the economics literature about experience curves and the economic historical literature about the massive WWII production effort.

There is large literature documenting the relation between increasing experience, defined as cumulative production, and decreasing costs.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

The literature has its roots in Wright's study of the aircraft industry but has now gone far beyond that, with studies ranging in aggregation level from individual plants to total global output of individual products.[2] It is clear that growing experience is positively correlated with cost declines, but the relative importance of the different channels through which this happens is an unsettled debate.

The most common explanation is simply learning-by-doing: individual workers and managers learn while they are on the job, improve the manufacturing process and thereby lower costs (Newell and Rosenbloom 1981; Heathcote, Brown, and Mewhort 2000). Military products have played an important role in the development of the concept of learning-by-doing in economics. According to Hirsch (1956), much of the early impetus to study the evolution of costs came from the U.S. Air Force, which was interested in cost decreases in airplane production. Theodore Wright himself was head of U.S. aircraft production during WWII and is reputed to have used his own law to forecast aircraft production costs. The Air Force commissioned a number of reports, particularly from RAND in the late 1940s and 1950s. These early reports suggested strong learning-by-doing effects in aircraft production (Asher 1956; Alchian 1963). Using WWII U.S. shipbuilding data on the well-known Liberty Ships, Rapping (1965) attributed the large decrease in person-hours[3] per unit to organizational and individual learning from production experience.

However, since the publication of these studies in the mid-twentieth century, scholars have extended the simple learning-by-doing interpretation, focusing on capital intensity, quality of management, knowledge spillovers, and R&D.

A higher capital intensity is a particularly important driver of the reduction in labor needed per unit of production (Scott-Kemmis and Bell 2010; David 1974). Thompson (2001) collected better data on capital intensity in Liberty ships production, and found that learning-by-doing was largely overestimated, with labor productivity improving mostly thanks to higher capital stock. Mishina (1999) studied the Boeing Seattle plant producing the B-17 Flying Fortress aircraft in detail and was able to control for capital intensity. He found that productivity stopped growing when production was scaled down, and he rejected cumulative

[2] Wright (1936), Rapping (1965), Sheshinski (1967), David (1974), Yelle (1979), Argote and Epple (1990), Thompson (2001, 2007), Nagy et al. (2013), and Lafond et al. (2018).

[3] Rapping (1965), as well as the primary literature from which we draw our raw data, uses the term manhours rather than person-hours. We use the term "person-hours," which is more historically accurate given the well-documented importance of women in WWII military production.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

production as a source of cost reduction. This is a finding that does not generalize in our dataset. Bahk and Gort (1993) worked with plant-level data, measuring experience as cumulative production but adding a variable for capital vintages. One of their conclusions is that when the capital vintages variable is included, the effect of calendar time disappears—or even reverses—suggesting that "industry-wide" learning is actually capital embodied.

Looking at the production of ships during WWII, Searle (1945) found that the rates of cost reduction varied from one yard to another more than they did between different types of vessels within a given yard. This lends support to the idea that the quality of management within an organization has a significant effect on the cost of production, as Argote and Epple (1990) have argued. In a classic paper about the cotton spinning industry, Lazonick and Brush (1985) found evidence that management-worker relations impacted work intensity and productivity improvements.

Another well-studied source of cost improvement is experienced in the production of another model or by another firm. Thornton and Thompson (2001) extended Thompson's (2001) dataset with 57 different types of ships and studied a variety of experience spillovers, both between and across yards. They found that such spillovers could be quite important, though this is sensitive to the indicator used. Intergenerational externalities (between different generations of a product) have been identified by Irwin and Klenow (1994) and by Benkard (2000). Irwin and Klenow (1994) also found strong international interfirm externalities in their study focused on the DRAM industry. Levitt, List, and Syverson (2013) estimated the effects of experience on labor productivity and defect rates in an automotive plant and documented in exceptional detail how learning by the workers of one shift was transmitted to workers of another shift.

More recent studies have focused on the effects of R&D on cost reduction. Sinclair, Klepper, and Cohen (2000) studied detailed product-level learning curves in the chemical industry, and found that the products for which there was the tightest connection between cost improvement and production experience were also those for which specific (process) R&D projects had been developed. They argue that because cumulative production conditions future returns to R&D, often cumulative production will be associated with R&D and thus to cost decrease. Funk and Magee (2015) have focused on the pre-commercial period of production. They showed that significant pre-commercial technological progress occurs before large-scale production experience occurs. Looking at firm-level, nonquality-adjusted unit costs in India, Dosi, Grazzi, and Mathew (2017) found that R&D-intensive sectors tended to exhibit a "negative"

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

learning rate, pointing toward R&D as a source of quality improvement, leading to higher costs rather than cost savings. Another explanation for heterogeneous learning rates is design complexity: if engineers perform improvements by trial-and-error on parts of a more complex entity, more complex designs require more coordinated improvement, which is less likely (McNerney et al. 2011).

Despite detailed datasets and attempts at uncovering sophisticated channels that affect productivity, many studies tend to suffer from a more basic issue: because experience is a cumulative variable and it tends to grow exponentially, the inclusion of a time trend to correct for omitted variables results in serious multicollinearity that prevents precise estimation of relative effects.

Studies where including a time trend led to unstable results include Papineau (2006), Söderholm and Sundqvist (2007), and landmark papers in the experience curve literature, such as Benkard (2000) and Bahk and Gort (1993). The Liberty ships studies face the same issue. Thompson (2001) does not control for a time trend, and Thornton and Thompson (2001) do include a time trend but find that it has "the wrong sign." In their regressions of defect rates on experience in an automotive plant, Levitt, List, and Syverson (2013) report opposite signs for the time trend and experience when using a small dataset (weekly data), but not with a larger (daily) dataset. This may be expected, as more data points compensate for the collinearity between the regressors.

A more complex issue is causality. While one approach has been to use instrumental variables (Benkard 2000; Söderholm and Sundqvist 2007, Witajewski-Baltvilks, Verdolini, and Tavoni 2015), we instead choose to exploit a classic natural experiment: military production during WWII, a setting in which we can gather comprehensive evidence, demand is exogenous, and the time trajectory of production is far from exponential, solving basic issues of causality and multicollinearity. However, it remains beyond the scope of this paper to settle the debates on which of the many channels of productivity improvement are most important. As noted by Levitt, List, and Syverson (2013), the relationship between productivity and cumulative production may be incidental to a deeper causal mechanism, involving direct accumulation of knowledge at the firm level, and possibly other sources of increasing efficiency when considering sector or product level experience curves.[4] Our contribution

[4] In a recent working paper, Ilzetzki (2022) takes an in-depth look at the data on aircraft plants (one of our three datasets), which he complements with newly digitized data on utilization rates to show that plants with higher utilization rates had higher improvement in productivity—they were "learning by necessity."

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

is to focus on one simple problem—the relative contribution of exogenous and endogenous factors to productivity growth—and to address it with as much data as possible. We stick to a simple distinction between causal factors that are well proxied by product-specific accumulated experience, and causal factors that are independent of this but are the same across all products.

A UNIQUE HISTORICAL CONTEXT

The demand shock to military equipment production in the United States caused by WWII was enormous. President Roosevelt made this clear in his "Arsenal of Democracy" fireside chat on 29 December 1940, when he said,

> I want to make it clear that it is the purpose of the nation to build now with all possible speed every machine, every arsenal, every factory that we need to manufacture our defense material (Roosevelt 1940).

When war broke out in Europe in September 1939, the U.S. Army had only 189,839 men and women (Herman 2012). By the time of Pearl Harbor, it was more than 1.4 million men and women strong (The National WWII Museum 2015). Defense spending increased almost 30 times between 1940 and 1942. The United States supplied the Soviet Union and the United Kingdom with war materiel, particularly the United Kingdom after enacting the Lend-Lease policy Act in 1941. Qualitative evidence suggests that the period between the autumn of 1939 and the American entry into WWII was important for preparing the United States for war production. This ramp-up continued throughout the war. Between 1943 and the end of the war in 1945, the U.S. government spent four times as much money on war production as it had spent in 1942 (Koistinen 2004).

However, these huge expenditures were not allocated to clear and consistent munitions orders. Instead, "munitions orders kept changing …as the project size of the American Armed Forces increased as the American military strategy evolved" (Carew 2009). As documented in the Army's own history of its role in economic mobilization in WWII, "automatic supply gave way in large measure to supply on the basis of specific requisitions from theatre commanders. This permitted procurement and issue of supplies more closely tailored to specific theatre needs as indicated by operation experience and changes in strategic plans" (Smith 1959, p. 173).

Support for battlefield victory at all costs continued throughout the war. From the highest office to the internal operations of the responsible

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

bureaucracies for military production and procurement, the priority was battlefield victory—rather than cost savings. In August 1943, the War Department Procurement Review Board was appointed to evaluate the composition, essentiality, and balance of procurement programs. It noted that "a war cannot be run like an industry; the criterion is not low costs but victory" (Smith 1959, p. 159). In 1940, the War Production Board set forth 12 criteria for the placement of contracts, including speed of delivery, quality, and price. In contrast, the War Production Board statement of March 1942 contained three ordered criteria: speed of delivery; conserving of superior facilities for the most difficult items of production; and placement of contracts with firms needing the least amount of additional machinery and equipment (Smith 1959, p. 263). This strongly suggests that companies were operating at maximum capacity and that using available capacity, rather than finding the cheapest supplier, was the key driver of procurement decisions. Price and cost efficiency criteria were reintroduced explicitly in late 1942 and 1943 but were never leading criteria.

The procedures for managing costs and awarding contracts evolved during the course of the war. For products whose structure and costs of production were well known, *Fixed-Price* contracts were the government's first preference. In situations where firms were creating new products with unknown costs the government instead used *Cost-Plus-a-Fixed-Fee* (*CPFF*) contracts. In 1941 contracts began to include renegotiation clauses, which meant that at the conclusion of the contract a renegotiation would take place between firms and the government to assess actual costs (Smith 1959). This practice was codified with the passage of the 1942 Renegotiation Act and linked legislation that made renegotiation a condition of all contracts or subcontracts greater than $100,000 (later revised up to $500,000). Producers were allowed to maintain a fluidly defined "fair profit" in renegotiation, which was typically between 5 and 13 percent. Renegotiation ultimately saved the government approximately $3 billion on 118,000 contracts (Koistinen 2004). Furthermore, it provided an incentive to conduct the detailed cost accounting that generated the data central to this paper.

While this was the general structure of contracting, procurement generally became more complex as a function of the complexity, novelty, and expense of the product being procured. Our comprehensive data captures this range of experiences because it expands the sample of products studied beyond just a handful of complex products. A clear example of this is the well-documented contracting process for aircraft during WWII (Holley 1964). For the government to reach deals with aircraft

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

and airplane engine makers, there was a longer initial negotiation process and then a more comprehensive renegotiation than for less complex products. This allowed both parties to adjust agreed-upon terms as they learned more about production costs and predicted efficiencies in production—including those predicted by Wright's law, which the government frequently used as an analytical tool in negotiation. In short, while there are commonalities across contracting for WWII, the process was not completely homogenous.

The universal inclusion of renegotiation clauses meant that contractors had limited incentives to decrease their costs of production, as they would only be compensated for final demonstrated costs. As has been already noted in the WWII learning literature (Bajari and Tadelis 2001), there is little incentive in a cost-plus contract to supply a product more cost-efficiently. That said, while the contracts themselves do not provide clear incentives for learning, the process of defense procurement does: a high-performing supplier could expect more orders from the government (Rogerson 1994). This possibility of repeat contracts as an incentive for learning and driving costs down may limit our exogeneity argument—at least when seen at the firm-level. However, the combination of the vast scale of the war effort and the intense pressure to ramp production up quickly made controlling costs difficult and required manufacturers to operate at near maximum capacity. The vast scale of the effort was so large that the majority of manufacturing firms in the United States were engaged in war-related production, and, together with the accelerated contracting process, this limited the ability of the government to select the most efficient producers when awarding contracts.

At the sectoral level, an argument can be made that cost played a role in determining which equipment was produced. The U.S. government intended to produce the deadliest weaponry at the lowest cost,[5] so that the *relative* demand for products that were substitutes was not independent of price. For instance, to the extent that two types of planes could be substituted for one another, the cheaper one may have been favored. Nonetheless, this was balanced by the fact that military strategy requires a diverse portfolio of different types of military equipment deployed in unison. The war could not be won by planes alone, but also required naval vessels, tanks, troop carriers, bullets, and a host of different types of equipment. Even within a category such as planes, military history suggests product complementarity rather than substitutability, for instance

[5] Rohlfs, Sullivan, and Kniesner (2016) use evidence of substitution between military personnel and weaponry to estimate the statistical value of life.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

between bombers and fighters.[6] In any case, our data is quite comprehensive, covering individual products (which may have been substitutes) and aggregates (that are clearly not substitutes).

While we focus on the American context of wartime production and the effects of this large demand stimulus, several papers suggest there may be other international data available for consideration and comparison. For example, Budrass, Scherner, and Streb (2010) study learning-by-doing and fixed-price contracts for aircraft in Germany during the war, while Okazaki (2011) focuses on Japanese airframe production. However, these papers—like earlier work on U.S. war production—focus on one narrow type of wartime production, in this case, aircraft. Furthermore, while the United States was able to produce for the war with relatively limited interruption from battles on the home front, the same was not true for Axis powers or even European Allies. As Tooze (2005) carefully demonstrates, German production was severely compromised by the Allied siege as the war progressed. Working far from the frontlines, the United States was able to ramp up production in an exceptional way and, as our data indicate, maintain extremely detailed records about wartime production.

DATA

Another key specific advantage of the historical context is that because many products were bought by the government, prices were recorded at a fairly detailed level and are now part of the public record.[7] Using a number of different sources, we created three datasets, as summarized in Table 1. The datasets are ordered from the most fine-grained to the most aggregated. The first dataset comprises time series of production and labor productivity for 152 types of aircraft (or plant-aircraft pairs), ships, and motor vehicles, 23 of which have never been published before. The second dataset consists of time series of production for 523 products from all categories of war materiel, paired with "early" and "late" unit costs. This dataset is new. The third dataset comprises time series for indices of contract prices for several branches of the War Department. While this last dataset was easy to collect, we are not aware of any study that has matched the relevant tables to construct experience curves from it.

[6] Indeed, Ilzetzki (2022) uses the relative demand for aggregate categories of aircraft as an instrument to estimate the causal impact of demand shocks on productivity in aircraft production at the level of plant-model pairs, and confirms a positive effect.

[7] The lack of availability of prices at the most granular level is a well-known problem for estimating productivity (Klette and Griliches 1996; Foster, Haltiwanger, and Syverson 2008).

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

TABLE 1
SUMMARY OF THE THREE DATASETS

Dataset	Sources	N	Time Span	Cost Data	Aggregation
Labor productivity	*Source Book*, Searle (1945), and Ford archives	152	01/1940 to 11/1945, $T \in [2,64]$	Person-hours per unit	Plant or product
OMPUS-USMH	USMH and OMPUS	523	08/1942 to 08/1945, $T = 2$	"Early" and "Late" "Standard Dollar Weight" per unit	Product
Contract prices	Crawford and Cook (1952)	10	01/1942 to 08/1945, $T = 44$	Index of contract prices	War (sub) departments

Notes: N is the number of time series and T is the number of observations in each time series. Time span refers to cost data. In each dataset, we have production data starting in January or July 1940.
Source: The authors.

We now describe each dataset more precisely, leaving the full details of data collection to Online Appendix A.

Labor productivity in Aircraft, Ships, and Motor Vehicles

Our first dataset, which we will refer to as *Labor productivity*, includes time series of labor productivity from three different sources (see Table 2). The first source is the well-known time series of labor productivity for ships, extracted from Searle (1945). It has been already widely extended and analyzed in Rapping (1965), Thompson (2001, 2007), and Argote, Beckman, and Epple (1990).

The second source is an extension of the data used in the study of aircraft by Alchian (1963), extracted from the source quoted in his paper, referred to as the *Source Book* (Army Air Forces 1947). It contains data on plant-product pairs, so the same plane in different plants is a separate time series, and so are different planes within the same plant.

The third source is entirely original. We have collected a novel set of time series from Ford's archive. This covers 23 products, mostly motor vehicles. Online Appendix A.1 further discusses data collection and transcription issues for each of these three sources.

Total Unit Costs at the Product Level

Our second dataset, which we will refer to as *OMPUS-USMH* or *OMPUS* for short, was created by combining two archival publications.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

TABLE 2
SOURCES OF THE *LABOR PRODUCTIVITY* DATA

Source	Type of Products	*N*	Aggregation Level
Searle (1945)	Ships	5	Product
Source book	Aircrafts	124	Plant-product pair
Ford's archives	Mostly motor vehicles	23	Product

Source: The authors.

The first publication, *Official Munitions Production of the United States, by Months, July 1, 1940–August 31, 1945 (OMPUS) Handbook* (Civilian Production Administration 1947), provides information about the monthly output of particular categories of munitions and specific products. The second publication is the *United States Munitions Handbook* (USMH), which provides the costs of production, in terms of "Standard Dollar Weights" at two moments—one at the beginning of the war and one towards the end—for many of the products recorded in the OMPUS. While the OMPUS is available online, the formerly classified USMH was located in the National Archives by one of us (D.G.).

Matching OMPUS to the USMH was relatively easy because each product listed in the USMH includes the page and column to which it should be matched in the OMPUS. One of the nice features of these data is that they are nearly comprehensive. The OMPUS publication was intended to record the "munitions figures of *all of the procurement services*" (our emphasis), and while the USMH does not report cost values for the majority of products, the coverage appears reasonably representative.

The main issue we had with the dataset concerned the dates at which the costs were recorded; the dates in the USMH are the same for every product category, but some products started production after the "early" date or ended production before the "late" date—in this case, we assumed, guided by the USMH explanatory notes, that the early/late dates corresponded to the start/end of production. Online Appendix A.2 provides a more detailed discussion of our interpretation of early and late costs and Standard Dollar Weights.

While the OMPUS provides production data, it does not provide data on experience, that is, cumulative production since the first unit. It is impossible to know what the previous manufacturing experience was for each product listed in the OMPUS. While other studies of experience curves have not corrected for prior experience or have applied a correction based on the assumption of exponential growth, we opted for an approach that corrects at the level of semi-aggregated categories, for

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

which we could obtain rough estimates or introduce one ourselves. For each product i, we estimated prior experience as a proportion of total war production experience. More precisely, we estimated initial experience as a category-specific factor ζ_i times the total amount of known production of product i during the war $\sum_{\tau=1}^{T} Q_{i,\tau}$. Denoting experience $Z_{i,t}$, we have

$$Z_{i,t} = \zeta_i \sum_{\tau=1}^{T} Q_{i,\tau} + \sum_{\tau=1}^{t} Q_{i,\tau}, \tag{1}$$

where T is the total number of months from January 1940 to August 1945. The first term represents estimated initial experience, and the second term is the usual experience computed as observed cumulative production up to time t.

Because these products are too detailed, we attempted to estimate the factors ζ_i at a more aggregated level, using the 81 most detailed levels of the table of contents of the OMPUS. The ζ_i and the extensive discussion of how we estimated them are provided in Online Appendix B. To give one example, consider fighter aircraft. We have 12 models of fighters in the USMH-OMPUS data, and we do not know the initial experience of each model. However, from other sources, we were able to estimate that around 300,000 planes were produced during WWII, and a total of 60,000 were produced at any time before the war. Therefore, we applied a value of $\zeta \approx 60{,}000/300{,}000 = 0.20$ to all aircraft models. For example, 680 units of the P–61 Black Widow model (OMPUS ref. 15/3) were produced during the war, so we constructed the experience variable as if $0.20 \times 680 \approx 136$ models had been produced before.

Contract prices at the War Department Level

Our last dataset, which we will refer to as *Contracts*, contains contract price indices and is taken from the statistics on procurement in Crawford and Cook (1952). This data allows us to construct experience curves for the period January 1942–August 1945 at several different levels of aggregation. The data is for the eight different branches of the War Department. The highest level of aggregation is the entire War Department. It is then subdivided into the Air Forces (AAF) and the Service Forces (ASF). The latter is then subdivided into seven technical services, which "were the operating agencies of the War Department in all its supply activities" (Smith 1959, p. 114).

The data on production, while expressed in millions of dollars, can be taken as indicative of production *volume* because it has been generated from a large sample of physical quantities, evaluated at 1945 prices rather

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

than at the prices in force during each month. The indices of contract price changes were computed from price changes to contracts for individual products, representing about half of the total value of War procurement. As for the OMPUS data, we estimated prior experience as a category-specific factor times the total level of wartime production. The detailed justification for the correction factors applied to this data is also given in Online Appendix B.

EMPIRICAL FRAMEWORK

Our objective is to estimate the effect of experience on cost. As discussed previously, simply regressing cost on experience, as is done in learning curve studies, suffers from several related issues, all of which are absent or limited in the context of WWII military production.

To restate the issues, because demand is generally elastic, the relationship between price (or cost) and production (or experience) embodies both demand elasticity and the causal effect of experience on cost. However, while demand relates price/cost and production, the effect of experience is between cost and *cumulative* production. It is common (outside of WWII) that production grows exponentially, so cumulative production grows exponentially too, at the same rate. Because of this, if we use production and cumulative production interchangeably, we can think of the issue as a problem of simultaneity between cost/price and production/experience (Nordhaus 2014), even though the mechanisms are in principle from experience to cost and from price to production.

Second, it is easy to argue that cost reductions cannot possibly come only from the growth of product-specific experience, whatever these proxies are.[8] The simplest possible way to acknowledge omitted variables is simply to allow for an exogenous exponential time trend, to capture overall economy-wide trends. However, as discussed in the previous paragraph, experience often grows exponentially and in a smooth manner (Lafond et al. 2018), making it hard to distinguish the effect of experience from the effect of an "exogenous" exponential time trend. Technically speaking, this is just an issue of imperfect multicollinearity.

[8] Thompson (2012) shows that if demand is elastic and cost decreases are purely driven by a trend but with autocorrelated noise, then it is likely that regressions of (log) costs on (log) cumulative production and a trend will find significant negative effects for cumulative production and negative effects for the time trend that are weaker than the truth. However, in the same setting, regressing the *change* in (log) costs on the change in (log) cumulative production results in an *opposite* bias, with an estimated effect of experience that is positive, and an effect of time (the intercept) that is more negative than the truth. Here, we systematically estimate models in terms of levels and first-difference.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

In the specifications in first-differences, where the exogenous time trend is simply the intercept of the regression, this is just an issue of limited variance of the regressor. While in principle this does not bias the coefficients, including both experience and a time trend usually makes standard errors extremely large and gives unstable results, as discussed in the Literature Review section. It is therefore a serious issue.

Nordhaus's Model

Nordhaus (2014) introduced a very simple model that illuminates the issues and helps clarify why demand exogeneity makes the identification of the effect of experience on costs possible.

Let Q_t denote production at time t and let Z_t denote cumulative production. Assume the unit cost function c_t depends on both experience Z_t and an exogenous time trend, with the functional form

$$c_t = c_0 Z_t^{-b} e^{-at}. \tag{2}$$

The parameters a and b are expected to be positive since we generally observe technological progress. Assume that production is equal to demand and price is equal to unit cost.[9] Furthermore, assume that demand has constant elasticity $\varepsilon > 0$ and that there is an exogenous and growing demand e^{dt}. This gives[10]

$$Q_t = D_t = D_0 c_t^{-\varepsilon} e^{dt}. \tag{3}$$

By letting Δ represent the time difference operator $\Delta x = x_t - x_{t-1}$, Equations (2) and (3) can be written as

$$\Delta \log c = -a - b\Delta \log Z \tag{4}$$

and

$$\Delta \log Q = -\varepsilon \Delta \log c + d. \tag{5}$$

[9] There could have been time varying markups during the war and heterogenous markups across producers. In the *Labor productivity* dataset, overall cost declines were very strong, so we believe that markup fluctuations would have had to be exceptionally strong to affect our results substantially. In the *OMPUS-USMH* and the *Contract prices* dataset, cost reductions are much smaller, so it is possible that cost decreases are over- or under-estimated due to unobserved markup fluctuations. Very little is known about market structure and pricing during the War. This is a valuable avenue for further research.

[10] Nordhaus (2014) also includes population growth in the exogenous demand term.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

Now, suppose that production grows exponentially. Then experience also grows exponentially at the same rate, that is

$$\Delta \log Q \approx \Delta \log Z.$$

Using this in Equation (5), the solution of the system of Equations (4) and (5) is

$$\Delta \log c = \frac{-a - bd}{1 - b\varepsilon}, \tag{6}$$

$$\Delta \log Z = \Delta \log Q = \frac{a\varepsilon + d}{1 - b\varepsilon}. \tag{7}$$

Nordhaus (2014)'s critique is that experience curve studies typically assume that the logarithm of experience is the single explanatory factor for the logarithm of costs. In other words, they assume that

$$\Delta \log c = \beta \Delta \log Z, \tag{8}$$

and perform a regression of the logarithm of costs against the logarithm of experience in order to measure the empirical parameter β. Inserting Equations (6) and (7) into Equation (8), we can write the parameter β as the ratio

$$\beta = \frac{\Delta \log c}{\Delta \log Z} = \frac{-a - bd}{a\varepsilon + d}. \tag{9}$$

It is clear that in general, the empirically estimated parameter β not only depends on the experience parameter $-b$, but it also depends on the exponential trend parameter a for costs, the exponential trend parameter d for exogenous demand growth, and the demand elasticity ε. The estimated β can be interpreted as being purely the effect of experience only when there is no exogenous time trend. In other words, if $a = 0$ then $\beta = -b$.

Let us now consider the case in which demand is completely exogenous ($\varepsilon = 0$), but for simplicity still growing exponentially. In this case Equations (6) and (7) become

$$\Delta \log c = -a - bd, \tag{10}$$

and

$$\Delta \log Z \approx \Delta \log Q = d, \tag{11}$$

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

so that putting Equation (11) into Equation (10) gives

$$\Delta \log c = -a - b\Delta \log Z, \tag{12}$$

which is the same as Equation (4). If demand is completely exogenous, we can estimate the supply equation directly.[11] A simple first-difference regression model for the logarithm of cost against the logarithm of experience with an additive constant is able to separate the effects of the exogenous trend parameter a and the "learning" parameter b.

While illustrative of the reverse causality issue, this model is not stochastic and all variables grow exponentially, so it cannot do justice to the other important issue, multicollinearity. If production grows exponentially with fluctuations, since computing experience from production involves integrating and therefore smoothing, Z will increase exponentially but with much lower fluctuations, so the term $\Delta \log Z_t$ will be approximately equal to the constant d, for all t.[12]

Fortunately, we will find empirically that during WWII, demand exogeneity led to production patterns that were *not* exponential, making the correlations between the regressors (time, production, and experience) much lower than in other studies using recent data.

Main Specification

It is customary in the literature to motivate the main specification for learning curve regressions by starting from the standard Cobb-Douglas production function. This does not assume exponentially increasing production, it allows for economies of scale, and it can be used with data on labor productivity.[13] With constant technical parameters and

[11] Pozzi and Schivardi (2016) used survey estimates of demand elasticity coupled with data on firm-level prices to separate the effects of demand and productivity on firm growth.

[12] To give a concrete example, if production follows a geometric random walk with drift r and noise σ, then the change in the log of cumulative production has a variance approximately equal to $\sigma^2 \tanh(r/2)$ (Lafond et al. 2018), which is much smaller σ^2. For instance, if we consider the Air Force time series from the *Contract prices* data, taking the period where we have monthly data and the growth is roughly exponential (January 1942 to January 1944), we have $\sigma = 0.047$ and $g = 0.068$. The measured variance of the change in the log of *cumulative* production is an order of magnitude less than σ, 0.0063, fairly close to the prediction from the formula of Lafond et al. (2018), 0.0087.

[13] In this model, optimal requirements for labor and capital grow at the same rate. This may not be an appropriate assumption, especially during WWII, when there was an important build-up of the capital stock (Gordon 1969). However, a more general framework allowing for factor-biased technological change would require more data for estimation, so we limit ourselves to the Cobb-Douglas case here.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

a time-varying Hicks-neutral productivity factor A_t that depends on experience and an exogenous time trend, as in Rapping (1965), we have

$$Q_t = A_t K_t^{\theta_k} L_t^{\theta_l},$$
$$A_t = Z_t^b e^{at}, \tag{13}$$

where K_t and L_t are capital and labor in period t and θ_k and θ_l are constants, with an additional term for an exponential exogenous productivity improvement as in Nordhaus (2014).

If firms minimize their total cost subject to the constraint of fixed *exogenous* demand $Q_t = \bar{Q}_t$, and factor prices are constant, the optimal labor demand per unit of production can be written

$$\log \frac{L_t^*}{\bar{Q}_t} \equiv \log l_t = B_L - \left(\frac{a}{s}\right) t - \left(\frac{b}{s}\right) \log Z_t + \left(\frac{1}{s} - 1\right) \log \bar{Q}_t, \tag{14}$$

where B_L is a constant and where $s = \theta_k + \theta_l$ represents economies of scale. Since we have data on the unit labor requirements $l_t = L_t^*/\bar{Q}_t$ in our *Labor productivity* dataset, we can estimate this equation. In the two other datasets, we have data on total unit costs $c_t = C_t/\bar{Q}_t$. Inserting optimal factor requirements in the total cost function $C_t = wL_t^* + rK_t^*$ gives

$$\log \frac{C_t}{\bar{Q}_t} \equiv \log c_t = B_C - \left(\frac{a}{s}\right) t - \left(\frac{b}{s}\right) \log Z_t + \left(\frac{1}{s} - 1\right) \log \bar{Q}_t, \tag{15}$$

where B_C is another constant. Apart from the constants, the Equations (14) and (15) are the same. They become exactly the same if we estimate a model in first differences, that is

$$\Delta \log c = \Delta \log l = -\left(\frac{a}{s}\right) - \left(\frac{b}{s}\right) \log Z_t + \left(\frac{1}{s} - 1\right) \log \bar{Q}_t. \tag{16}$$

Our favored specification is to estimate Equation (16) in a pooled panel. However, each dataset presents specific patterns of missing data and thus different opportunities for robustness checks and alternative specifications—typically fixed effects for the levels (Equations (14) and (15)) and separate regressions for each time series. Thus, we present our detailed econometric specifications separately for each dataset.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

EMPIRICAL RESULTS

We now present results from performing regressions on the three data sets here.[14] A point of prior optimism comes from examining the correlations between the independent variables. As shown in Table 6, they range from about 50 to 90 percent. Although this might seem high, they should be compared with systematically higher correlations in other contexts, such as the chemical industry numbers reported by Lieberman (1984), which are all greater than 90 percent.

Labor productivity in Aircraft, Ships, and Motor Vehicles

A typical model specification (Lieberman 1984) for analyzing a panel of experience curve time series is the fixed effects model

$$\log l_{it} = \kappa_i + \alpha\, t + \beta \log Z_{it} + \gamma \log Q_{it} + \eta_{it}. \tag{17}$$

Comparing to Equation (14), $\kappa_i = B_L$, $\alpha = -a/s$, $\beta = -b/s$, and $\gamma = 1/s - 1$. The results of this model applied to the *Labor productivity* dataset are reported[15] in Table 3. Including only experience, we find that $\beta \approx -0.33$, which implies a drop of cost of 20 percent for each doubling of experience. This number coincides with the original observation by Wright and is seen in many other studies of products such as airplanes. The estimated coefficient for the exogenous time trend is not very statistically significant, but its value is relatively high; given that the data is monthly, a negative coefficient of 0.004–0.006 implies an annual rate of exogenous technological progress of about 5 to 7 percent. Finally, including production indicates very mild and not very significant economies of scale. Throughout the paper, we will find no convincing evidence of non-constant returns to scale.[16]

A typical feature of experience curves regressions is that they have highly autocorrelated residuals (Thompson 2012). We ran the test of Wooldridge (2002) for AR(1) residuals in short panels and found in all three cases very strong evidence against the null of no first-order autocorrelation. This suggests estimating a model with autocorrelated disturbances, or a model with lags of the dependent variable, which we do in Online Appendix D.1.

[14] The replication files are available from Lafond, Greenwald, and Farmer (2022).

[15] Throughout the paper, we report significance at the 5, 1, and 0.1 percent levels. Standard errors are clustered by product. For fixed effect models, we report the "within" R^2; for first-difference models with no intercept (no "time" variable), we do not report the R^2 because it cannot be interpreted as the explained proportion of variance.

[16] The number of observations is lower because we removed months during which production was null.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

TABLE 3
PANEL REGRESSION RESULTS FOR *LABOR PRODUCTIVITY*

	Fixed Effects			First Differences		
Experience	–0.326***	–0.304***	–0.276***	–0.253***	–0.217***	–0.203***
	(0.017)	(0.020)	(0.024)	(0.020)	(0.022)	(0.025)
Time		–0.004	–0.006		–0.022***	–0.024***
		(0.003)	(0.003)		(0.004)	(0.005)
Production			–0.036*			–0.010
			(0.015)			(0.007)
N	3,034	3,034	2,981	2,830	2,830	2,740
R^2	0.75	0.75	0.75		0.13	0.13

Source: The authors.

But since we are not interested in the intercepts, and since from a theoretical point of view, we expect that it is the change in experience that induces a change in cost, we favor a first-differences specification,[17] that is

$$\Delta \log l_{it} = \alpha + \beta \, \Delta \log Z_{it} + \gamma \, \Delta \log Q_{it} + \eta_{it}. \tag{18}$$

The role of experience is fairly consistent with the fixed effects regression, but it indicates a larger role in the time trend. The coefficient β now ranges from –0.25 when only experience is included to –0.20 when all independent variables are included and remains solidly statistically significant. However, the coefficient α for the exogenous time trend now has a value of –0.024 when all variables are included, corresponding to a much larger annual improvement rate of about 29 percent, and is now strongly statistically significant. As we will discuss in the Discussion section, this represents a third of the rate of cost decrease, with the other two-thirds accounted for by experience. When including production, the coefficient γ is now very small and statistically insignificant.

Finally, we present the results of independent regressions for each individual product or plant, removing all time series with less than 10 available cost values (so 118 of 152 time series remain). Since the coefficient for production is so small in the aggregate regression, we exclude the production variable here in the interest of parsimony. The equation estimated is

$$\Delta \log l_{it} = \alpha_i + \beta_i \, \Delta \log Z_{it} + \eta_{it}. \tag{19}$$

[17] Unfortunately, our panel is strongly unbalanced and contains missing observations, limiting our options for unit root and cointegration tests, but see Online Appendix D.1. Also note that since the dataset contains missing values, we lose more observations than the number of products (152) when taking first differences.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

Figure 2 compares the estimated coefficient $\hat{\beta}$ for the regression Equation (19) when the exogenous time trend is included to the case where it is not, that is, we restrict $\alpha_i = 0$ and Equation (19) becomes a regression through the origin. We observe a great deal of heterogeneity in experience rates, but the results are broadly consistent with the aggregate regression. The variation is a mixture of intrinsic differences in real progress rates and statistical fluctuations due to the fact that the individual regressions are based on short time series (we have also checked that there is no correlation between the estimated slope parameter and the number of available observations). With some exceptions, most of the data is clustered along the identity line, demonstrating that the coefficient estimated for experience is not overly sensitive to the restriction $\alpha = 0$. This stands in sharp contrast with the results reported by Nordhaus (2014). Figure 2 in Nordhaus (2014), which is analogous to our Figure 2, shows coefficients that are often extremely far from the "reasonable" range $\beta \in (0,–0.5)$, and which change by more than an order of magnitude when a time trend is included. We believe that the main difference is that in his data, many of the production time series (and thus experience time series) are close to following an exponential trend, creating extreme problems of multicollinearity.

In summary, for the *Labor productivity* dataset, we observe a strong and clear relationship between declining labor costs and experience, with exponents $\beta \approx –0.2$. We also show a significant role for the exogenous time trend, whose overall effect is comparable to that of experience (the Discussion section will provide an estimate of the relative importance of each variable for all the datasets).

OMPUS-USMH: Unit Costs at the Product Level

We now turn to the *OMPUS* dataset, which gives total unit costs instead of labor productivity and contains a more complete sampling of military products. A limitation is that we observe the cost for each product on only two different dates, and these dates differ across products. Nonetheless, because we assume that the parameters are the same for all products, we can exploit cross-sectional heterogeneity. We do this in two ways.

First, we are still able to estimate a fixed-effects model, but we also interpret it as a "differences" estimator. As is well-known, when $T = 2$, the first-differences and the fixed-effects estimators are the same. Online Appendix C shows that this result holds when the differences are heterogenous, in the sense that the span of time between the two observations is different from one product to another. We perform inference and report R^2 from the fixed-effects estimation.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

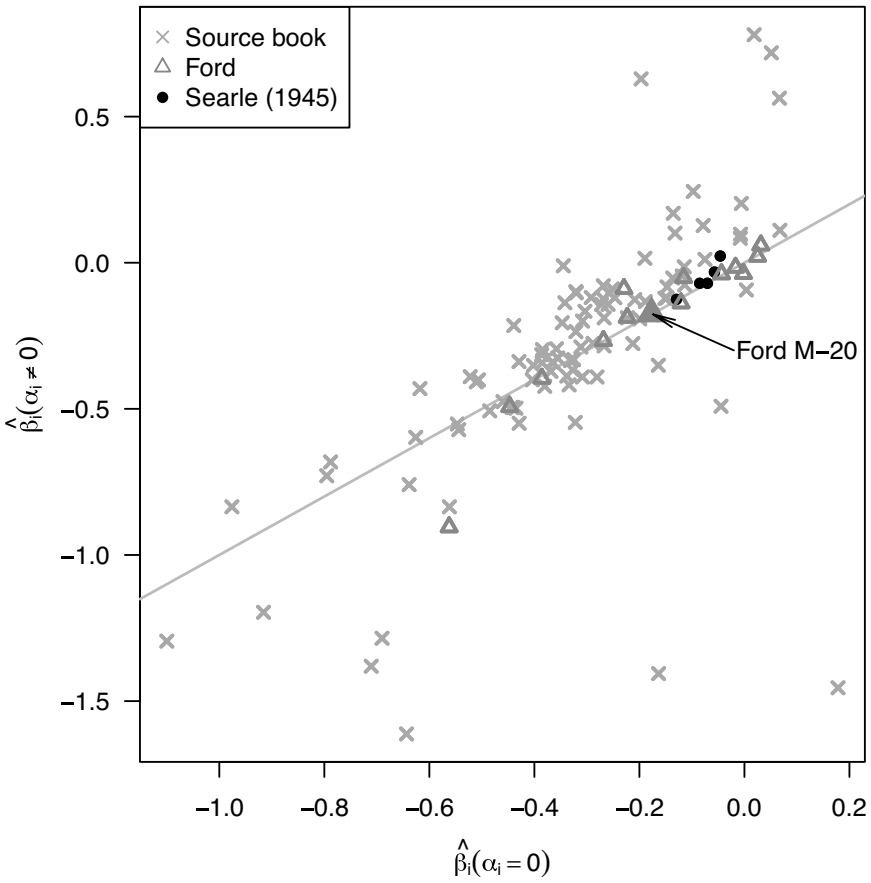


FIGURE 2
COEFFICIENT β_i FOR EXPERIENCE WHEN THE TIME TREND IS INCLUDED, PLOTTED AGAINST THE SAME COEFFICIENT WHEN THE TIME TREND IS NOT INCLUDED, USING EQUATION (19)

Notes: The grey line is the identity. Three values of β_i, which were greater than two in absolute value have been excluded. The Ford Armored Car discussed in the introduction (Figure 1) is shown as a slightly larger solid triangle.
Source: Own estimates based on the *Labour Productivity* dataset.

Table 4 shows the results of the regression, first including only experience and then including an additional time trend and production. The coefficients for experience are statistically significant but are smaller than for the *Labor productivity* time series, ranging from –10 percent when the other variables are not included to about –5.5 percent when time is included. The coefficients for the exogenous time trend are also significant, and about 0.4 percent per month or 5 percent per year (about 60 percent of the rate of cost decreases). As before, the coefficient for production is small and statistically insignificant.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

TABLE 4
PANEL REGRESSION RESULTS FOR *OMPUS-USMH*

	Fixed-Effects/Heterog. Differences			Growth Rates Cross Section		
Experience	–0.098***	–0.055**	–0.058**	–0.086**	–0.079	–0.098
	(0.015)	(0.017)	(0.019)	(0.031)	(0.040)	(0.052)
Time		–0.004***	–0.005***		–0.002	–0.002
		(0.001)	(0.001)		(0.003)	(0.003)
Production			0.008			0.024
			(0.009)			(0.044)
N	1,046	1,046	1,002	523	523	482
R^2	0.13	0.17	0.19		0.06	0.06

Source: The authors.

Second, since with two observations we can compute an average growth rate, we perform a cross-sectional regression of the growth rates (Columns (4)–(6); see Online Appendix C, Equation (25) for details). The results are similar, although with a sensibly higher coefficient for experience and lower statistical significance.

The upshot is that the results for the *OMPUS-USMH* dataset are broadly consistent with those of the *Labor productivity* dataset, albeit with smaller experience exponents.

Contract prices at the War Department Level

Figure 3 shows costs against experience in our last dataset on contract price indices at the War Department level. "Total" (on the right) represents the entire War Department, which is divided into Army Air Forces (AAF) and Army Service Forces (ASF); the latter is subdivided into the other seven services. Apart from Quartermaster, which saw its price *increase* over the period,[18] the other services have a relatively similar slope.

In Figure 4, we show the experience coefficient obtained from separate regressions for each individual time series, with and without the exogenous time trend. We make four observations that will be useful as a background to motivate the panel specification and understand its result.

First, in most cases, we see that standard errors are higher when the time trend is included, again reflecting the multicollinearity problem. Second, for some subservices (Ordnance, Signal, Transport, and Engineers), the coefficient is only slightly affected by the inclusion of the time trend, which in this case confirms that the lack of correlation allows

[18] Quartermaster contains common goods such as clothing, which were produced extensively prior to the war.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

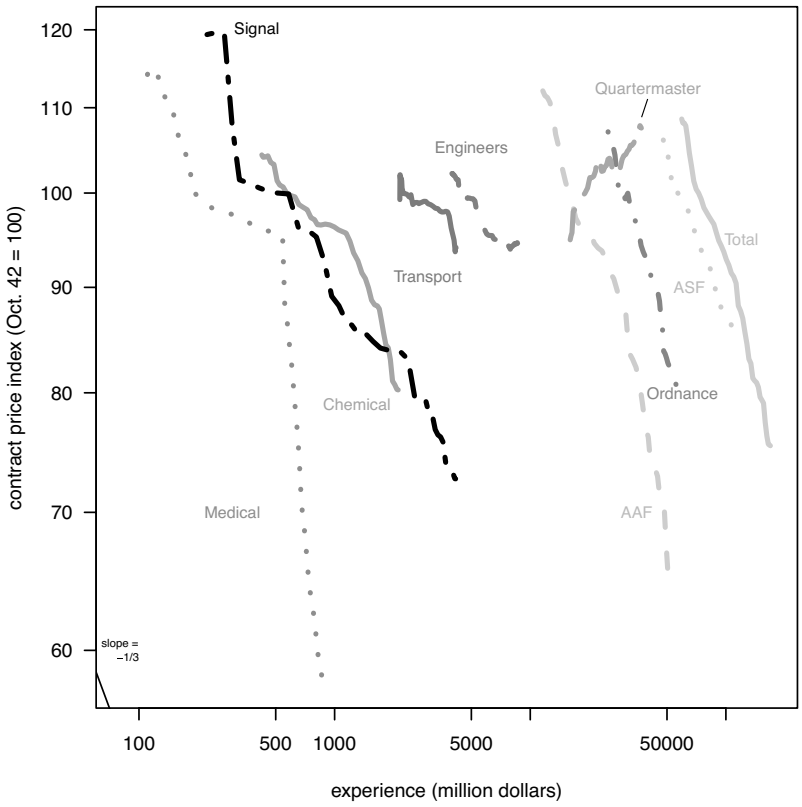


FIGURE 3
EXPERIENCE CURVES FOR THE WAR DEPARTMENTS AND COMPONENT AGENCIES

Source: Crawford and Cook (1952).

us to pin down the relative effects of the exogenous trend and experience. However, for some other subservices (Medical, Chemical), the estimated coefficient changes so much that its sign flips. Third, when the time trend is included, all subservices have confidence intervals overlapping zero, indicating a lack of statistical significance. Fourth, the confidence intervals of all subservices largely overlap, suggesting that a panel regression assuming a unique slope could give sensible results.

Because Quartermaster has such an obviously different slope in Figure 3 (on the top right), we remove it from the dataset and present panel regression results for seven time series (AAF and six subservices of ASF).

Table 5 shows the results. In line with previous estimates, we find an economically significant effect for both experience and the time trend, and a weak effect for production.[19]

[19] Note that statistical significance levels are based on hypothesis testing using the Student distribution with $G - 1$ degrees of freedom, where G is the number of clusters, which here equals 7.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

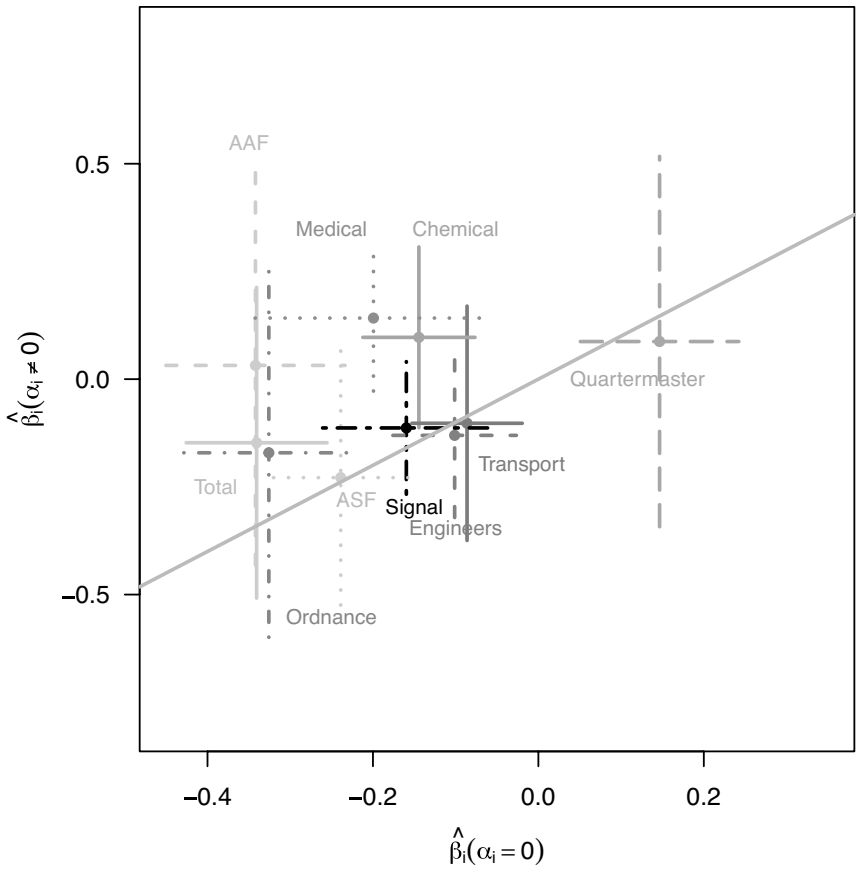


FIGURE 4
ESTIMATED COEFFICIENT OF THE FIRST DIFFERENCE REGRESSION OF THE LOG OF CONTRACT PRICES ON THE LOG OF EXPERIENCE, INCLUDING OR EXCLUDING AN EXOGENOUS TIME TREND

Notes: The lines show plus or minus 2 HAC standard errors. The solid grey line is the identity.
Source: Authors' estimates based on the *Contract prices* dataset.

DISCUSSION

In this section, we take stock of the results obtained from the three datasets and briefly summarize our robustness checks.

Summary of the Results

To provide an overview of our results, we have gathered some key statistics in Table 6. The first three columns show the average growth rates of the variables, computed as the average one period change in their logarithm, pooling all observations for which both $\Delta \log c$ and $\Delta \log Z$ are

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

TABLE 5
PANEL REGRESSION RESULTS FOR *CONTRACTS*

	Fixed Effects			First Differences		
Experience	–0.206**	–0.150*	–0.165**	–0.189***	–0.106	–0.118*
	(0.037)	(0.051)	(0.044)	(0.024)	(0.045)	(0.040)
Time		–0.003	–0.003		–0.004	–0.004
		(0.003)	(0.002)		(0.003)	(0.003)
Production			0.040*			0.006*
			(0.016)			(0.002)
N	308	308	308	301	301	301
R^2	0.77	0.78	0.81		0.05	0.06

Source: The authors.

available.[20] The next three columns show the Pearson correlations on pairwise complete observations between the "within" transformation of the regressors. The next two columns report the coefficients for the main regression specification, which includes experience (coefficient β) and time (coefficient α) and is estimated in first-differences (fixed-effects or equivalently "heterogenous differences" for the *OMPUS*, see Online Appendix C). The last column shows the ratio of α to the average growth rate of cost, which is the share of cost decrease accounted for by the exogenous trend.[21]

As is evident from the first column of the table, costs in the *Labor productivity* dataset, which are unit labor requirements, decline more than seven times faster than either of the two unit costs. One explanation is simply that improvement rates are heterogenous and the *Labor productivity* data comes overwhelmingly from the rapidly improving aircraft industry, while the other datasets on unit costs have a much wider and balanced coverage. Another explanation could be that capital deepening during the period improved labor productivity more than overall productivity.

The three columns giving correlations between the independent variables demonstrate that multicollinearity is a limited issue in our data. Although many of these correlations are high, they are still much lower than those reported in most previous studies (see, for instance, Lieberman 1984, table 2).

The following two columns provide an overview of our regression results. In every case, both experience and the exogenous time trend

[20] For the *OMPUS*, only two observations per product are available, at the starting and ending points t_0 and t_1, so the growth rates shown are averages over all products of the monthly growth rates $\left(\log x_{t_1} - \log x_{t_0}\right)/\left(t_1 - t_0\right)$. In the *Labor productivity* dataset, in contrast to other datasets, the growth rate for experience is much larger than the growth rate of production, partly because in this dataset, we did not correct for initial experience.

[21] For the *OMPUS*, this is computed based on the fixed effects equation; see Online Appendix C.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

TABLE 6
DESCRIPTIVE STATISTICS AND MAIN RESULTS

	Growth Rates			Correlations			Coefficients		Share Exo.
	c	Z	Q	Q,t	Q,Z	Z,t	β	α	$\alpha/g.r.(c)$
Labor productivity	–6.6	20.1	4.4	46.9	78.2	80.8	–21.7	–2.2	33.3
OMPUS-USMH	–0.8	7.4	5.2	72.3	70.5	59.2	–5.5	–0.4	60.4
Contract prices	–0.8	3.5	3.0	62.9	56.5	85.7	–10.6	–0.4	54.3

Notes: All values are multiplied by 100. Growth rates are monthly. $g.r.(c)$ is the growth rate of costs.
Source: The authors.

play an important role. Given the larger rate at which costs decline in the *Labor productivity* dataset, it is not surprising that both effects are substantially larger in this dataset.

In the last column, we divide the exogenous trend rate α by the overall rate of cost decline (the figures in the first column of the table). This ratio is roughly a third for the *Labor productivity* dataset but 60 percent for the *OMPUS-USMH* and a half for the *Contract prices* datasets. Online Appendix D9 shows that while the estimated coefficients α and β are likely to be different across datasets (considering the standard errors associated with these estimates), the hypothesis that cost decreases are due to both endogenous and exogenous processes is relatively similar amount is much more defensible.

The overall conclusion is that we see a substantial effect of both experience and an exogenous time trend, in roughly equal proportions. This makes it clear that technological improvement is neither fully exogenous nor fully endogenous. Once we control for time and experience, the effect of monthly output is negligible. When multicollinearity is not as severe as in other settings, and enough data are pooled together, the estimates are relatively stable across specifications.

Robustness of the Results

We discuss the robustness of the results to changes in econometric methodology, assumptions about initial experience, and omitted variables. The details of the robustness checks, when possible, are in Online Appendix D.

Time series analysis (D.1).

Overall, we tend to find similar results to those reported in the text if we (a) include time fixed effects instead of an exponential trend, (b) replace experience by its lagged value, and (c) test for times series properties and estimate the models suggested by these tests.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

Heterogenous coefficients (D.2).

We allowed for fixed effects in the first difference regression, thereby allowing for time trends to differ. This had no effect on the *Labor productivity* dataset, but for the *Contracts* data, the effect of experience disappears completely, although standard errors are large enough that we cannot reject a value such as the one estimated in the main text. Estimating Swamy's (1970) or Pesaran and Smith's (1995) heterogenous coefficient models gives the same outcome, with robust results for *Labor productivity* but not for *Contracts*.

Instrumenting by lagged values (D.3).

Our dependent variable is a total (either person-hours or cost) divided by output. Thus, output appears on both sides of the equation. We instrumented output and experience by their lagged value and found that this does not change the results.

Alternative assumptions on initial experience (D.4).

We re-estimated our models using different assumptions about initial experience. Overall, it is clear that changing the levels of initial experience changes the results, often in the direction of attributing a larger role to the exogenous time trend. Nonetheless, under a very wide range of possible corrections, the signs remain negative. Without implausible changes to our estimates of initial experience, such as a factor of 5, the share of the exogenous time trend cannot exceed about 70 percent.

Increasing quality and design changes.

Technological progress may increase the quality of a product rather than reduce its cost. This is a limited issue here, as many qualitative accounts of war production describe it as mass production of fixed designs, although not without tensions.[22] For instance, Best (2018) describes how mass production techniques were implemented for producing aircraft, which then were sent to Air Force depot and modification centers for fitting the weapons, which are more susceptible to design changes. These overall limited design changes suggest that our study may not suffer too

[22] Design changes requirements were cited by Ford as one of the reasons for the delays in reaching full speed production of the B-24 Liberator plane at the famous purpose-built large-scale Willow Run plant in Detroit (Baime 2014). At General Motors, Keizer had to implement "flexible mass production" techniques (Herman 2012). Mishina (1999) reports eight versions of the B-17 "Flying Fortress" since 1934, but notes that the design was frozen at model B-17E, after only 134 of a total of 6,981 B-17s had been built. Changes between the B-17E and the B-17 F&G were small changes intended to accommodate the changing fitted weapons, which themselves featured continuous and important upgrading.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

much from missing data on quality but also somewhat limit the external validity of our results, as many experience curves relevant to current-day policies include the production of successive designs.

Research and development.

A common criticism of experience curves is that they miss a key driver of cost reduction, R&D. We were not able to find any monthly product-level R&D data. R&D, and especially federal R&D, was important during the war, grew very fast, and ultimately impacted patenting and the spatial distribution of employment and innovation (Gross and Sampat 2020). For instance, the modification centers mentioned earlier were themselves linked to the Office for Scientific Research and Development and the National Advisory Committee for Aeronautics (ex NASA), the major institutions organizing war-related R&D. However, our estimates of *both* α and β are very likely to be biased in the same direction because one would expect R&D to depend positively on both time and experience. Thus, it is unclear a priori if including R&D would change our results on the relative effects of endogenous and exogenous progress. However, because we are silent on the role of R&D or ultimate causal factors more generally, our study is unable to provide a more fine-grained policy recommendation.

Comparing USMH and Contracts (D.5).

In principle, the raw data for these two datasets largely overlaps, so they should give similar results if the *USMH* data can be aggregated at the same level as the *Contracts* data. We used our own concordance table to test whether estimating the time and experience parameters at the War Department level in each dataset would give similar results. We found only moderate agreement, but the standard errors are large, so we cannot exclude that this is due to a low sample size.

Controlling for inflation (D.6).

An obvious factor in determining the cost is the price of inputs, which we are unable to observe for each product. We used the producer price index for all commodities as a control, as a single regressor or interacted with individual dummies in an attempt to acknowledge that different products have different input mixes. Our main results do not change.

External learning (D.7).

We attempted to estimate spillover effects by adding the growth of aggregate production experience (total monthly war spending) as a regressor. We did this for both production and cumulative production.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

The effects of these variables are not consistent across datasets, and this leaves our estimates for the effects of our own experience unchanged.

Depreciation of experience (D.8).

The literature has emphasized that part of past experience may become irrelevant. We constructed experience variables using a retention factor δ, that is $Z_t = \delta Z_{t-1} + Q_t$. We found that allowing for $\delta < 1$ tends to increase the share of the cost decrease due to exogenous factors. However, even when this improves the fit (mostly for *Labor productivity*), the improvement is not large and the share of exogenous progress remains relatively close to our main estimates.

Assuming known economies of scale (D.10).

Because it is difficult, due to collinearity, to simultaneously estimate economies of scale and experience effects, one can assume a known value for economies of scale and check how the estimated effects of experience and exogenous progress change. In general, assuming slightly decreasing economies of scale gives a slightly better fit to the data and leads to a smaller estimated contribution of exogenous progress.

CONCLUSION

This paper takes advantage of new and wide-ranging data to fill a gap in the existing economic historical literature between macro-histories of the conflict and its effects on national economies and the narrowly focused learning-by-doing studies that examine single product lines. In focusing on this in-between scale of inquiry, this paper has also shown that military production during WWII offers a context that solves key issues with experience curves: endogeneity, and multicollinearity between time, production, and experience. Because the demand for weapons was driven by battlefield needs, which grew and then shrank, not only was it exogenous, but production, experience, and time were relatively uncorrelated. As a result, we found that, in contrast to other studies, estimating the partial effect of experience on cost was relatively robust to the inclusion of a time trend. Despite remaining issues of data quality, we conclude that both experience and other effects captured by an exogenous time trend were quantitatively important during the War, in roughly equal proportion. This resolves an important issue raised by Nordhaus (2014) on the difficulties of calibrating the relative amounts of endogenous and exogenous progress in integrated assessment models of climate change. Of course, the WWII context was very specific, and our study suffers

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

from clear issues of external validity. But it still provides, in our opinion, a useful datapoint that could help current modelers make more informed choices on the relative role of exogenous and endogenous progress.

A point that deserves some emphasis is that the effect of experience on cost differs among products. As a result, when assessing a policy intervention for a given product, one must first determine whether there is historical evidence that increasing experience is associated with decreasing cost. With this caveat about heterogeneity in mind, our study broadly confirms that investing in a specific product in order to stimulate experience will cause production costs to decrease. However, our study also finds that a large share of technological progress cannot be attributed to product-specific experience. This suggests that to make targeted investments in a particular technology fully effective, policies should also aim at fostering the broader innovation system.

Overall, while taking a step back to evaluate the simple dichotomy between exogenous and endogenous progress in the WWII context was necessary, we hope that some of our data could be reused for more detailed historical studies of the causes of technological and industrial progress in specific industries, perhaps highlighting the role of market structure and knowledge diffusion more than what we were able to do while committing to using all available cost and production data.

REFERENCES

Alchian, Armen. "Reliability of Progress Curves in Airframe Production." *Econometrica* 31, no. 4 (1963): 679–93.

Argote, Linda, Sara L. Beckman, and Dennis Epple. "The Persistence and Transfer of Learning in Industrial Settings." *Management Science* 36, no. 2 (1990): 140–54.

Argote, Linda, and Dennis Epple. "Learning Curves in Manufacturing." *Science* 247, no. 4945 (1990): 920–24.

Army Air Forces. *The Source Book of World War II Basic Data. Airframe Industry, Vol. I, Direct Man-Hours - Progress Curves*. Washington, DC: Office of Air Force, 1947.

Asher, Harold. *Cost-Quantity Relationships in the Airframe Industry*. Santa Monica, CA: RAND Corporation, 1956. Available at https://www.rand.org/pubs/reports/R291.html.

Bahk, Byong-Hyong, and Michael Gort. "Decomposing Learning by Doing in New Plants." *Journal of Political Economy* 101, no. 4 (1993): 561–83.

Baime, Albert J. *The Arsenal of Democracy: FDR, Detroit, and an Epic Quest to Arm an America at War*. New York: Houghton Mifflin Harcourt, 2014.

Bajari, Patrick, and Steven Tadelis. "Incentives Versus Transaction Costs: A Theory of Procurement Contracts." *RAND Journal of Economics* 32, no. 3 (2001): 387–407.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

Benkard, C. Lanier. "Learning and Forgetting: The Dynamics of Aircraft Production." *American Economic Review* 90, no. 4 (2000): 1034–54.

Best, Michael H. *How Growth Really Happens: The Making of Economic Miracles Through Production, Governance, and Skills*. Princeton, NJ: Princeton University Press, 2018.

Budrass, Lutz, Jonas Scherner, and Jochen Streb. "Fixed-Price Contracts, Learning, and Outsourcing: Explaining the Continuous Growth of Output and Labour Productivity in the German Aircraft Industry during the Second World War 1." *Economic History Review* 63, no. 1 (2010): 107–36.

Carew, Michael G. *Becoming the Arsenal: The American Industrial Mobilization for World War II, 1938–1942*. Lanham, MD: University Press of America, 2009.

Civilian Production Administration. *Official Munitions Production of the United States*. Washington, DC: CPA, 1947.

Crawford, Richard H., and F. Cook Lindsey. *US Army in World War II: Statistics – Procurement*. Fort Leavenworth, KA: Archives Section–Library Services, 1952. Available at http://www.alternatewars.com/BBOW/Stats/USA_in_WW2_Stats-Procure_9-APR-52.PDF.

David, Paul A. "The 'Horndal Effect' in Lowell, 1834–1856: A Short-Run Learning Curve for Integrated Cotton Textile Mills." *Explorations in Economic History* 10, no. 2 (1974): 131–50.

Dosi, Giovanni, Marco Grazzi, and Nanditha Mathew. "The Cost-Quantity Relations and the Diverse Patterns of 'Learning by Doing': Evidence from India." *Research Policy* 46, no. 10 (2017): 1873–86.

Fishback, Price V. "The Second World War in America: Spending, Deficits, Multipliers, and Sacrifice." In *The Economics of the Second World War: Eighty Years On*, 53–57. VoxEU, ebook, 2020.

Foster, Lucia, John Haltiwanger, and Chad Syverson. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98, no. 1 (2008): 394–425.

Funk, Jeffrey L., and Christopher L. Magee. "Rapid Improvements with No Commercial Production: How Do the Improvements Occur?" *Research Policy* 44, no. 3 (2015): 777–88.

Gordon, Robert J. "$45 Billion of US Private Investment Has Been Mislaid." *American Economic Review* 59, no. 3 (1969): 221–38.

Gross, Daniel P., and Bhaven N. Sampat. "Inventing the Endless Frontier: The Effects of the World War II Research Effort on Post-War Innovation." NBER Working Paper No. 27375, Cambridge, MA, 2020.

Hall, Robert E. "By How Much Does GDP Rise If the Government Buys More Output?" *Brookings Papers on Economic Activity*, no. 2 (2009): 183–231.

Heathcote, Andrew, Scott Brown, and D.J.K. Mewhort. "The Power Law Repealed: The Case for an Exponential Law of Practice." *Psychonomic Bulletin & Review* 7, no. 2 (2000): 185–207.

Herman, Arthur. *Freedom's Forge: How American Business Produced Victory in World War II*. New York: Random House, 2012.

Hirsch, Werner Z. "Firm Progress Ratios." *Econometrica* 24, no. 2 (1956): 136–43.

Holley, Irving Brinton. *Buying Aircraft: Materiel Procurement for the Army Air Forces*. Washington, DC: Office of the Chief of Military History, Department of the Army, 1964.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

Ilzetzki, Ethan. "Learning by Necessity: Government Demand, Capacity Constraints, and Productivity Growth." LSE Working Paper, London, UK, March 2022.

Irwin, Douglas A., and Peter J. Klenow. "Learning-by-Doing Spillovers in the Semiconductor Industry." *Journal of Political Economy* 102, no. 6 (1994): 1200–27.

Klette, Tor Jakob, and Zvi Griliches. "The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous." *Journal of Applied Econometrics* 11, no. 4 (1996): 343–61.

Koistinen, Paul A. C. *Arsenal of World War II: The Political Economy of American Warfare, 1940–45*. Lawrence: University Press of Kansas, 2004.

Lafond, François, Aimee Gotway Bailey, Jan David Bakker, Dylan Rebois, Rubina Zadourian, Patrick McSharry, and J. Doyne Farmer. "How Well Do Experience Curves Predict Technological Progress? A Method for Making Distributional Forecasts." *Technological Forecasting and Social Change* 128 (2018): 104–17.

Lafond, François, Diana Greenwald, and J. Doyne Farmer. "Can Stimulating Demand Drive Costs Down? World War II as a Natural Experiment." Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2022-05-21. https://doi.org/10.3886/E170901V1.

Lane, Nathaniel. "The New Empirics of Industrial Policy." *Journal of Industry, Competition and Trade* 20, no. 2 (2020): 209–34.

Lazonick, William, and Thomas Brush. "The 'Horndal Effect' in Early US Manufacturing." *Explorations in Economic History* 22, no. 1 (1985): 53–96.

Levitt, Steven D., John A. List, and Chad Syverson. "Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant." *Journal of Political Economy* 121, no. 4 (2013): 643–81.

Lieberman, Marvin B. "The Learning Curve and Pricing in the Chemical Processing Industries." *The RAND Journal of Economics* 15, no. 2 (1984): 213–28.

McNerney, James, J. Doyne Farmer, Sidney Redner, and Jessika E. Trancik. "Role of Design Complexity in Technology Improvement." *Proceedings of the National Academy of Sciences* 108, no. 22 (2011): 9008–13.

Mishina, Kazuhiro. "Learning by New Experiences: Revisiting the Flying Fortress Learning Curve." In *Learning by Doing in Markets, Firms, and Countries*, edited by Naomi R. Lamoreaux, Daniel M. G. Raff, and Peter Temin, 145–84. Chicago: University of Chicago Press, 1999.

Nagy, Béla, J. Doyne Farmer, Quan M. Bui, and Jessika E. Trancik. "Statistical Basis for Predicting Technological Progress." *PloS One* 8, no. 2 (2013): e52669.

Newell, Allen, and Paul S. Rosenbloom. "Mechanisms of Skill Acquisition and the Law of Practice." In *Cognitive Skills and Their Acquisition*, edited by John R. Anderson, 1–55. New York: Psychology Press, 1981.

Nordhaus, William D. "The Perils of the Learning Model for Modeling Endogenous Technological Change." *Energy Journal* 35, no. 1 (2014): 1–13.

Okazaki, Tetsuji. "The Supplier Network and Aircraft Production in Wartime Japan." *Economic History Review* 64, no. 3 (2011): 973–94.

Papineau, Maya. "An Economic Perspective on Experience Curves and Dynamic Economies in Renewable Energy Technologies." *Energy Policy* 34, no. 4 (2006): 422–32.

Pesaran, M. Hashem, and Ronald Smith. "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics* 68, no. 1 (1995): 79–113.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

Pozzi, Andrea, and Fabiano Schivardi. "Demand or Productivity: What Determines Firm Growth?" *RAND Journal of Economics* 47, no. 3 (2016): 608–30.

Ramey, Valerie A. "Identifying Government Spending Shocks: It's All in the Timing." *Quarterly Journal of Economics* 126, no. 1 (2011): 1–50.

Rapping, Leonard. "Learning and World War II Production Functions." *Review of Economics and Statistics* 47, no. 1 (1965): 81–86.

Rogerson, William P. "Economic Incentives and the Defense Procurement Process." *Journal of Economic Perspectives* 8, no. 4 (1994): 65–90.

Rohlfs, Chris, Ryan Sullivan, and Thomas J. Kniesner. "Reducing Risks in Wartime Through Capital-Labor Substitution: Evidence from World War II." *Journal of Risk and Uncertainty* 52, no. 2 (2016): 163–90.

Roosevelt, Franklin D. "Fireside Chat 16. On the 'Arsenal of Democracy'." Miller Center of Public Affairs, University of Virginia, Presidential Speech Archive 29, Charlottesville, VA, 1940.

Scott-Kemmis, Don, and Martin Bell. "The Mythology of Learning-by-Doing in World War II Airframe and Ship Production." *International Journal of Technological Learning, Innovation and Development* 3, no. 1 (2010): 1–35.

Searle, Allan D. "Productivity of Labor and Industry." *Monthly Labor Review* 61, no. 6 (1945): 1132–47.

Sheshinski, Eytan. "Tests of the 'Learning by Doing' Hypothesis." *Review of Economics and Statistics* 49 (1967): 568–78.

Sinclair, Gavin, Steven Klepper, and Wesley Cohen. "What's Experience Got to Do with It? Sources of Cost Reduction in a Large Specialty Chemicals Producer." *Management Science* 46, no. 1 (2000): 28–45.

Smith, Ralph Elberton. *The Army and Economic Mobilization*. Washington, DC: GPO, 1959.

Söderholm, Patrik, and Thomas Sundqvist. "Empirical Challenges in the Use of Learning Curves for Assessing the Economic Prospects of Renewable Energy Technologies." *Renewable Energy* 32, no. 15 (2007): 2559–78.

Swamy, P. A. "Efficient Inference in a Random Coefficient Regression Model." *Econometrica* 38, no. 2 (1970): 311–23.

The National World War II Museum. *By the Numbers: The US Military*. New Orleans: The National World War II Museum, 2015. Available at http://www.nationalww2museum.org/learn/education/for-students/ww2-history/ww2-by-the-numbers/us-military.html.

Thompson, Peter. "How Much Did the Liberty Shipbuilders Learn? New Evidence for an Old Case Study." *Journal of Political Economy* 109, no. 1 (2001): 103–37.

———. "How Much Did the Liberty Shipbuilders Forget?" *Management Science* 53, no. 6 (2007): 908–18.

———. "The Relationship Between Unit Cost and Cumulative Quantity and the Evidence for Organizational Learning-by-Doing." *Journal of Economic Perspectives* 26, no. 3 (2012): 203–24.

Thornton, Rebecca Achee, and Peter Thompson. "Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding." *American Economic Review* 91, no. 5 (2001): 1350–68.

Tooze, J. Adam. "No Room for Miracles. German Industrial Output in World War II Reassessed." *Geschichte und Gesellschaft* 31, no. 3 (2005): 439–64.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press

———. *The Wages of Destruction: The Making and Breaking of the Nazi Economy*. London and New York: Allen Lane, 2006.

War Production Board. *Wartime Production Achievements and the Reconversion Outlook*, Report of the Chairman, Washington, DC, 1945.

Witajewski-Baltvilks, Jan, Elena Verdolini, and Massimo Tavoni. "Bending the Learning Curve." *Energy Economics* 52 (2015): S86–99.

Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

Wright, Theodore P. "Factors Affecting the Cost of Airplanes." *Journal of the Aeronautical Sciences* 3, no. 4 (1936): 122–28.

Yelle, Louis E. "The Learning Curve: Historical Review and Comprehensive Survey." *Decision Sciences* 10, no. 2 (1979): 302–28.

https://doi.org/10.1017/S0022050722000249 Published online by Cambridge University Press