Universal Child Care, Maternal Employment, and Children's Long-Run Outcomes: Evidence from the US Lanham Act of 1940

Chris M. Herbst, *Arizona State University and IZA*

This paper analyzes the US Lanham Act of 1940, a heavily subsidized and universal child care program administered during World War II. I first estimate its impact on maternal employment using a triple-differences model. I find that employment increased substantially following the introduction of the program. I then study children's long-run labor market outcomes. Using Census data from 1970 to 1990, I assess well-being in a life-cycle framework by tracking cohorts of treated individuals throughout their prime working years. Results from difference-in-differences models suggest the program had persistent positive effects, with the largest benefits accruing to the most economically disadvantaged adults.

> In what is today a nearly forgotten social experiment, the federal government subsidized nationwide child care for working mothers of young children during World War II. It was the first time in the nation's history that day care for children who were not poor was supported by public funds. (Geraldine Youcha, *Minding the Children*, 1995, 307)

I. Introduction

The US Federal Government, in response to the surge in women's employment, administered a system of near-universal child care throughout

I am grateful to seminar participants at the University of Arizona's Department of Sociology and the University College London-University of Bergen Labor and Child Care Workshop, as well as conference participants at the Association for

[*Journal of Labor Economics*, 2017, vol. 35, no. 2]

Submitted October 29, 2014; Accepted December 10, 2015; Electronically published January 31, 2017

World War II. Popularly known as the Lanham Act of 1940, the child care program was considered a temporary war emergency measure and was aimed at providing children ages 0–12 with a safe environment so that mothers could contribute to the nation's war production effort.[1] Federal Lanham Act grants were awarded to communities based on a demonstrated need for war-time child care; at its apex, the program was administered in over 635 communities in every state except New Mexico (Stoltzfus 2000). Although it operated for only a brief period—from 1943 to 1946—the Lanham Act dispensed over $1 billion (in 2012 dollars) to construct and maintain child care facilities, to train and pay teachers, and to provide meal services.

The Lanham Act is widely considered a milestone in the history of US child care policy, primarily because it was the first, and only, federally administered program to serve children regardless of family income. Yet there is virtually no understanding of the program's implications for a number of policy-relevant outcomes, including maternal employment and long-run child well-being. Such evidence is increasingly important, given the interest in the United States and elsewhere in creating universal early care and education programs. Although several US jurisdictions offer universal preschool programs—including Georgia, Oklahoma, and Boston—they have not been operating long enough to shed light on whether the positive short-run impacts found in recent evaluations translate into long-run schooling and labor market success (Gormley and Gayer 2005; Fitzpatrick 2008; Cascio and Schanzenbach 2013; Weiland and Yoshikawa 2013). Conversely, evidence on the long-run impact of early childhood programs comes almost exclusively from highly targeted interventions, including the Head Start, Perry Preschool, and Abecedarian programs (e.g., Garces, Thomas, and Currie 2002; Ludwig and Miller 2007; Anderson 2008; Deming 2009). To my knowledge, the only evidence on the long-run impact of universal child care comes from a recent evaluation of the Norwegian system (Havnes and Mogstad 2011b). The goal of this paper, therefore, is to provide the first comprehensive analysis of a broadly accessible, heavily subsidized child care program in the United States context.

The paper begins by estimating the impact of the Lanham Act on maternal employment. To do so, I combine 1940 and 1950 Decennial Census data

Public Policy Analysis and Management and the Southern Economic Association for helpful suggestions. I particularly thank Joanna Lucio, Erdal Tekin, Melissa Kearney, Magne Mostad, Tarjei Havnes, Gordon Dahl, Liz Washbrook, Lane Kenworthy, Natalie Fousekis, and Resul Cesar for their detailed comments. Chen-Yu Kao, Andrew Kao, Zoram Kaul, and Casey Boyd-Swan provided excellent research assistance. Contact the author at chris.herbst@asu.edu. Information concerning access to the data used in this article is available as supplementary material online.

[1] The Lanham Act funded a variety of war-essential public works projects in addition to providing child care services. For brevity's sake, I refer to the child care program as the Lanham Act throughout the paper.

with a triple differences (DDD) approach to study labor supply at the participation and hours-of-work margins. The identification strategy exploits the large cross-state variation in the generosity of Lanham Act expenditures, comparing the effect of increased child care funding across women with and without age-eligible children after reform versus before reform. I find that the Lanham Act generated sizable increases in maternal employment. The baseline DDD estimate implies that increasing states' Lanham Act spending from the bottom to the top quartile of the distribution—equivalent to raising spending from $22 to $66 per child—would produce a 4.4 percentage point increase in the employment rate for treated women. This estimate is approximately one-quarter of the treated group's pre-reform mean. In addition, a supplementary analysis using 1940 and 1960 Census data finds that these employment effects persisted 17 years after the child care program was terminated.

I then turn my attention to studying the impact of the Lanham Act on children's long-run outcomes related to labor market and program participation. I utilize Decennial Census data between 1970 and 1990 to identify a set of treated cohorts, who were born between 1931 and 1946, as well as a set of comparison cohorts, who were born between 1947 and 1951. The use of three Census data sets enables me to trace the full life-cycle effects of the Lanham Act, starting when treated cohorts were in their mid-20s. I identify the impact of the child care program in a difference-in-differences (DD) framework, essentially comparing the difference in adult outcomes between treated and comparison cohorts born in states where Lanham Act spending was high with the difference for these groups where spending was low. I find that the Lanham Act had sizable positive effects on a variety of outcomes that persisted throughout adulthood. In particular, the baseline DD estimates imply that increasing states' child care spending from the bottom to the top quartile of the distribution would generate a 0.01 standard deviation increase in a summary index of five labor market and program participation outcomes. An auxiliary analysis of the Lanham Act's distributional effects reveals that the benefits of the child care program accrued largely to the most economically disadvantaged adults.

As with all DD designs, the analysis of adult outcomes must overcome two threats to its internal validity: (i) the presence of unobserved, contemporaneous shocks that differentially affected the outcomes of treated and comparison cohorts and (ii) the possibility that the outcomes of individuals born in low- and high-spending states would have trended differently in the absence of the Lanham Act. The key concern regarding the first threat is the shock represented by the war, which catalyzed a large exodus of men from the labor market and produced potentially important changes in children's home environment. I attempt to control for any war-induced changes in labor markets and home life by adding the state-level Armed Forces mobilization rate (defined as the fraction of men ages 18–44 who were drafted or enlisted for war) to the baseline DD model. This variable has been used by others to proxy the outflow of men from the labor market and the increased pressure on women to leave home for paid work (Acemoglu, Autor, and Lyle 2004; Goldin and Olivetti 2013).

To test for differential outcome trends, I exploit the fact that the comparison cohorts are drawn from a (post-reform) period in which the Lanham Act no longer operated. Given that the Lanham Act was terminated in 1946, I use the comparison cohorts—who were born in the first 5 years after the program ended—to test whether there are similar outcome trends for individuals born in low- and high-spending states. The tests provide evidence that the common trends assumption is satisfied. I also conduct a supplementary matched case control analysis, in which each high-spending state is paired with one or more low-spending states based on the degree of similarity in the outcome trends. The DD estimates continue to show positive long-run effects of the Lanham Act. Finally, I experiment with a placebo reform that keeps the Lanham Act "turned on" for some cohorts in the comparison group. The effect of the placebo is small in magnitude and never statistically significant.

Together these findings contribute to well-established literatures exploring the impact of early childhood programs on parental employment and child outcomes. In addition, this paper complements a body of work assessing the role of World War II in increasing women's labor supply. Although early work by Goldin (1991) suggests that the war was not a "watershed" event for American women, more recent papers by Acemoglu et al. (2004) and Goldin and Olivetti (2013) challenge this notion, finding that the war did in fact catalyze a long-run employment response. These latter papers rely on DD designs that exploit the measure of Armed Forces mobilization described above. To capture the causal effect of World War II, mobilization is assumed to be uncorrelated with other contemporaneous shocks. However, the introduction of universal child care represents one potential shock. Indeed, the historical record shows that the Lanham Act not only played a crucial role in the nation's war production effort but also fundamentally altered women's views on paid work and institutional child care. Such evidence is consistent with this paper's empirical finding that the program increased mothers' employment in the short run and the long run. Thus, it is plausible that the watershed event represented by World War II was aided by another watershed, the United States's first, and only, universal child care program.

The remainder of the paper proceeds as follows. The next section summarizes the relevant research on the impact of child care prices and programs on maternal employment and long-run child outcomes. Section III provides a detailed description of the Lanham Act. In particular, it summarizes the history and major design features of the program, provides insight into the characteristics of Lanham Act centers, and introduces the measure of Lanham Act spending used in the empirical analyses. Section IV is devoted to the analysis of maternal employment, while Section V implements the analysis of long-run outcomes. Finally, Section VI concludes.

II. Relevant Child Care Literature

A. Maternal Employment

Evidence on the relationship between child care programs and maternal work decisions comes from three sources: reduced-form and structural studies of child care prices, reduced-form studies of child care subsidy programs, and quasi-experimental evaluations of policy reforms. The most common methodological approach to estimating price effects includes a discrete choice participation equation with selection-corrected predicted hourly child care expenditures and wages as the key right-hand-side variables. Results from these studies consistently point to a negative relationship between child care costs and mothers' employment (Anderson and Levine 2000; Connelly and Kimmel 2003; Tekin 2007a; Herbst 2010). However, the range of estimated own-price elasticities is quite large: 0.06 to -1.36. The second approach examines the impact of actual childcare subsidy receipt on maternal employment. The most frequently studied program is the United States's Child Care and Development Fund (CCDF), an employment-based subsidy system targeting low-income families. The empirical framework models the employment decision using subsidy receipt as the key explanatory variable. The evidence suggests that CCDF-funded child care subsidies have large positive effects on the employment of economically disadvantaged single mothers (Tekin 2005, 2007b; Blau and Tekin 2007).

The final set of studies exploits geographic and temporal variation in the introduction of policy reforms (DD designs) or birthday-based discontinuities in program exposure (IV or RD designs). Regarding DD studies, Baker, Gruber, and Milligan (2008) and Lefebvre and Merrigan (2008) evaluate the introduction of universal child care in Quebec, Canada, while Havnes and Mogstad (2011a) examine Norway's universal program. United States–based DD studies include Cascio and Schanzenbach's (2013) analysis of universal pre-kindergarten programs in Oklahoma and Georgia, as well as Cascio's (2009b) paper on the introduction of kindergartens throughout the 1960s and 1970s. Regarding the latter design, Gelbach (2002) and Fitzpatrick (2012) use children's quarter-of-birth in the 1980 and 2000 Censuses, respectively, to instrument for kindergarten participation, while Fitzpatrick (2010) uses the discontinuity created by age-eligibility cut-offs to analyze the Georgia and Oklahoma pre-kindergarten programs. With the exception of the Canadian child care program, these interventions generate small employment effects that are applicable to specific subgroups.

B. Long-Run Child Outcomes

Most long-run evidence comes from studies of targeted education-focused programs or small-scale interventions. For example, Ludwig and Miller's (2007) analysis of Head Start exploits the discontinuity in local program allocations, finding positive effects on mortality and educational attainment.

Deming's (2009) paper, based on sibling fixed effects, finds that Head Start participants scored 0.23 standard deviations higher on a summary index of eight young adult outcomes. Positive long-run effects have also been estimated for participants in the Infant Health and Development, Perry, and Abecedarian programs (McCormick et al. 2006; McCormick et al. 2006; Anderson 2008; Heckman et al. 2010; Conti, Heckman, and Pinto 2015).

Although there is a large body of evidence on publicly funded child care programs, most of these studies are targeted at important subpopulations or are limited to analyses of short- and medium-run outcomes. For example, Baker et al. (2008) provide evidence on the short-run developmental effects of universal child care in Quebec, Canada, while Gupta and Simonsen (2010) and Black et al. (2014) provide medium-run evidence for the Danish and Norwegian systems, respectively. In the United States, Herbst and Tekin (2012, 2016) estimate the short- and medium-run impact of the CCDF, a subsidy program restricted to the working poor, while a large number of studies examine the short- and medium-run impact of pre-kindergarten programs, only some of which are universal (e.g., Gormley and Gayer 2005; Fitzpatrick 2008; Barnett et al. 2013; Cascio and Schanzenbach 2013; Weiland and Yoshikawa 2013). In addition, a noteworthy paper by Cascio (2009a), which estimates the long-run impact of kindergarten introduction in the United States, finds some positive effects for white adults but not for black adults.

To my knowledge, the only evidence on the long-run impact of universal child care comes from Havnes and Mogstad's (2011b) study of the Norwegian system. The reform was phased in between 1976 and 1979, expanding child care access for those ages 3–6. Exploiting the differential growth in child care coverage across municipalities, the author's DD estimates imply sizable increases in educational attainment, reductions in welfare participation, and delayed childbearing for treated cohorts in their early 30s. The results also show that most of the benefits accrued to individuals from economically disadvantaged backgrounds.

III. Background on the Lanham Act

A. History and Description

Prior to World War II, the Works Project Administration (WPA) operated a number of Depression-era child care centers that provided jobs for unemployed teachers.[2] The centers targeted low-income families, with the goal of increasing parental employment. With the United States' entry into the war—and men's subsequent exit from the labor market—it became clear

[2] This discussion draws on numerous historical accounts of the Lanham Act. For more information, see Koshuk (1947), Dratch (1974), Tank (1980), Anderson (1981), Riley (1994), Tuttle (1995), Youcha (1995), Stoltzfus (2000), and Fousekis (2011).

that large numbers of women were needed to bolster war production. Aided by the Rosie the Riveter campaign, the federal government urged women to join the war effort. However, it became clear that the stock of WPA centers would be insufficient for absorbing the growing demand for child care. Indeed, stories of children locked in cars adjacent to factories, chained to temporary trailer homes, and left in movie theaters quickly filled newspapers and eventually became the subject of congressional hearings (US Senate 1943a).

In response, federal funding for child care was approved in August 1942 through the National Defense Housing Act of 1940, also known as the Lanham Act. The Lanham Act was not designed to administer a system of child care. It was intended to fund the construction and maintenance of infrastructure projects deemed critical to the war effort, as well as community hospitals and schools.[3] Nevertheless, the Federal Works Agency (FWA) was assigned responsibility for distributing child care funds despite protests from child and education agencies, including the Children's Bureau and Office of Education. With the FWA in control, the program's intent was clear: to administer a system of temporary child care as a war expedient, not as a permanent expansion to the welfare state.

Federal funds for child care were available to construction and maintain facilities, to train and pay teachers, and to handle all other operating expenses. To access Lanham Act grants, communities in "war impact areas" had to show the FWA that they had insufficient resources to meet the surging child care demand.[4] Federal guidelines stipulated that communities had to engage in three activities to measure and demonstrate need: hold conferences on women's wartime work and child care; establish committees composed of parents, local nonprofits, and government agencies; and conduct needs assessment surveys. In the meantime, state legislatures created committees to document via public hearings local infrastructure and social service needs. The public hearings became the chief mechanism through which parents and advocates lobbied for child care centers, and results from the needs assessments were presented to state lawmakers. While local actors prepared applications for federal Lanham Act funds, state policy makers

[3] An interesting historical aside: the Lanham Act ultimately funded child care because the phrase "public services" in the original legislation was reinterpreted to include child care provision. It occurred by way of administrative decision by agencies in the executive branch, in particular the Federal Works Agency. It occurred without presidential action, congressional debate, or an alteration of the original legislative language. This led some to characterize Lanham Act child care funding as an "inspired afterthought" (Kerr 1973, 163).

[4] The term "war impact area" was a designation given to communities that were directly involved in the war production effort. For example, these communities contained defense contractors, industries manufacturing war-essential goods (e.g., textiles), and those involved in agricultural production.

passed legislation authorizing the establishment of child care centers, which would be run, in most cases, by the Department of Education. The review of applications was generally conducted at multiple bureaucratic levels, including the local WPA and FWA offices, a regional FWA office, and finally by federal administrators. Grants awarded by the FWA were matched by community contributions that were initially set at 50% of a project's total cost. However, in practice, widespread resource constraints limited the local contribution to one-third of all expenses. Most local funding came from parent copayments, initially set at $0.50 per child per day and rising to $0.75 in 1945.[5]

The first Lanham Act grants were distributed in July 1943. Although the program was scheduled to be terminated after Japan's surrender in September 1945, its funding continued until mid-1946 after protests from families and advocates claimed that child care services were needed until servicemen returned home from the war. Eligibility for Lanham Act subsidies was generally restricted to children ages 0–12, although administrators were initially reluctant to serve infants and toddlers. Children ages 0–5 were served in nursery environments, while those ages 6–12 were provided with before- and after-school care. Although there may have been a parental work requirement, it was largely de facto as opposed to explicit. Families were eligible regardless of income level, and the consumer subsidy was generous. At the war's peak, parents were required to contribute $0.75 ($9.50 in 2012 dollars) per child per day.

B. Quality in Lanham Act Child Care Centers

Lanham Act centers operated for long hours. The evidence suggests that they provided child care services 6 days per week, during most holidays, and throughout the summer.[6] It was common for preschool-aged children to spend at least 12 hours per day in the center, usually on a 6 a.m. to 6 p.m. schedule; school-aged children spent less time in care, typically for a few hours before and after school. Certified schoolteachers were among the first to be employed in the child care centers. To prepare nonteachers, cities generally contracted with universities to establish formal training programs. For example, the city of San Francisco, in collaboration with San Francisco State College, operated a well-known preparatory "academy" called the Teacher-Training Nursery School. In addition, the federal government cre-

[5] Federal and local funds were deployed for different purposes. Money from the former was used to build and maintain child care centers, train and pay workers, purchase all supplies, and cover the cost of all other operating expenses. Money from the latter was used to purchase food that was served to children.

[6] Yet the centers were also responsive to local needs. For example, some centers provided night-time and evening care to accommodate factories and airfields operating on 24-hour schedules, while others maintained seasonal programs.

ated a recommended 10–12-week training course aimed at volunteers who would have direct contact with children.

Anecdotal evidence suggests that preschool-aged children engaged in indoor and outdoor play; used educational materials such paints, clay, and musical instruments; and took regular naps. Children were provided with hot lunches, a snack, and dinner if necessary. The recommended child-teacher ratio in Lanham Act centers was 10:1, and many centers abided by the recommendation. Programs for school-aged children included breakfast and dinner, participation in music and drama clubs, library reading, and assistance with schoolwork. California centers were among the highest in quality: they had an explicit nutrition focus, children were given a medical exam, parents completed a developmental history, and teachers were provided with in-service training and college credit. On the other hand, low quality was pervasive in other areas. For example, a child care center in Baltimore reportedly contained 80 children in one room (with as many bathrooms), prepared meals on a hot plate, and required children to cross a highway in order to reach the playground. Thus, variation in quality across states and localities was likely to be substantial.

There were few systematic evaluations of children and parents using Lanham Act centers. Perhaps the best evidence comes from a descriptive analysis of children attending two centers in Bellflower, California, located in Los Angeles County (Koshuk 1947). The study examined administrative records on 500 children and parents from 1944 to 1946. Results from this analysis are illuminating. Perhaps the biggest concern from critics of the Lanham Act was that long hours in institutional care might fray the mother-child relationship. However, parent reports of children's behavior and family relationships suggest that such concerns were premature. Upon departure from the center, only 5.2% of mothers reported that the child was less willing to "obey or cooperate with adults," 0.6% claimed that the child became "less affectionate," and 1.7% felt that family relations were "less close." In addition, analyses of the teacher observations indicate that children made reasonable progress in specific developmental domains. For example, over 80% of children made "excellent" or "good" mental progress, relative to their status at entry, and about 75% of children made "excellent" or "good" social progress. Finally, despite parents' initial skepticism about institutional care, fully 100% of mothers reported that the "child enjoyed nursery school," and 81% had a "generally favorable" opinion of "early childhood education."

C. Measurement of the Lanham Act

To estimate the impact of the Lanham Act on maternal employment and long-run outcomes, I create a measure of total state-level Lanham Act spending per child ages 0–12. I create this measure by compiling annual child care expenditure data from a variety of congressional committee hearing reports, as well as from the publication *Annual Report, Federal Works Agency* (US Senate 1943a, 1943b; US House of Representatives 1945; Federal Works Agency, various years). These sources provide expenditure data for the 1943 calendar year and the 1945 and 1946 fiscal years. Therefore, my spending variable covers nearly the full period in which the Lanham Act operated. To create the denominator of the spending ratio, I estimate the population of white, non-Hispanic children (ages 0–12) in each state using the Integrated Public Use Microdata Series (IPUMS) of the 1940 US Decennial Census (Ruggles et al. 2015). The spending variable is adjusted to reflect constant 2012 dollars.

Lanham Act spending in the average state totaled about $58 per child ages 0–12 (median: $39), with a minimum of $0 in New Mexico and a maximum of $264 in California. The top five states are California ($264), Washington ($251), Oregon ($162), Florida ($141), and Arizona ($140). The bottom five states include Pennsylvania ($8), North Dakota ($6), West Virginia ($5), Idaho ($3), and New Mexico ($0).[7] Figure 1 provides a state-by-state look at Lanham Act spending. Specifically, I assign each state to a spending quartile and then color-plot the quartiles. A few observations are noteworthy. First, there is substantial variation in the geographic location of low- and high-spending states. For example, those in the top-half of the distribution include low-population Western and Southeastern states, as well as high-population states in the Northeast. Second, there is a reasonable amount of (political) ideological variation across the low- and high-spending states. Eighteen of the 38 states carried by Roosevelt in the 1940 presidential election are in the bottom half of the spending distribution, and three of the 10 states voting for Wendell Willkie—the losing candidate—are in the top half of the distribution. Thus, it appears that Roosevelt neither rewarded states voting for him nor penalized states voting against him. Finally, Lanham Act spending by coastal states was consistently greater than their inland counterparts: 8 of the top 10 states are on the coast. This is consistent with the location of most war production centers, as defined by the number of "war impact areas" in each state, because manufacturing firms near the coast could quickly move goods into overseas war theaters (US Senate 1943a).[8]

Table 1 explores more formally the state-level determinants of Lanham Act spending by estimating a series of regressions of child care expenditures on several state characteristics. The first four columns test several geographic and institutional variables. Column 1, which includes a binary indicator for coastal states, confirms that such states spent significantly more than noncoastal states. Column 2 includes a binary indicator for states voting

[7] Alaska and Hawaii are excluded from all descriptive analyses because they were not yet states.

[8] Coastal states contained about 17 areas, on average, designated as "war impact areas," compared to 11 in noncoastal states.

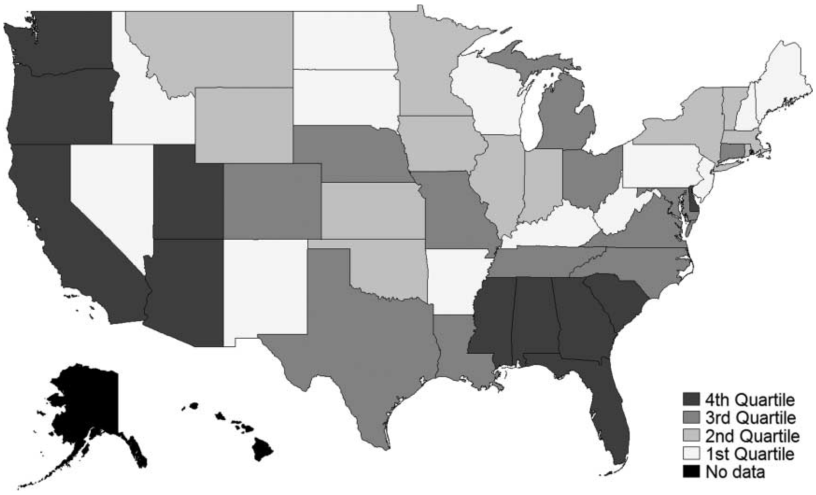

FIG. 1.—State-level Lanham Act child care expenditures (per child ages 0–12)

for Willkie in the 1940 Presidential election. The coefficient is statistically insignificant, suggesting again that funding decisions were not politically motivated. Column 3 adds a control for the Armed Forces mobilization rate, defined as the proportion of men ages 18–44 in each state who were drafted or enlisted for war.[9] Its coefficient is also statistically insignificant, indicating that the demographic forces shaping war enlistments are independent of the forces driving demand for war-time child care.

The last three columns test a few demographic characteristics. These variables are calculated using the 1940 Decennial Census from the IPUMS (Ruggles et al. 2015). To maximize war production, it is plausible that Lanham Act funds were distributed in a manner that favored states with higher levels of female educational attainment and employment. Columns 5 and 6 examine this by adding the average number of years of completed schooling by women and the female employment rate, respectively.[10] The coefficient on both variables is positive and statistically significant, providing evidence in favor of skill-biased funding decisions. When both are included in the regression with the coastal state indicator, as shown in column 7,

[9] Data on state-level mobilization rates can be found here: http://economics.mit.edu/faculty/dautor/data/autacemly06. Alternatively, they are included in Table A.1 of a working paper version of Goldin and Olivetti's (2013) published paper. The paper can be accessed here: http://www.nber.org/papers/w18676.pdf?new_window=1.

[10] The measure of female educational attainment is the one used in Acemoglu et al. (2004), and it can be accessed here: http://economics.mit.edu/faculty/dautor/data/autacemly06. All other measures were created by the author from the 1940 Census.

Table 1
State-Level Determinants of Lanham Act Child Care Spending

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coastal state (= 1)	45.61**			47.26**			52.97**
	(17.62)			(19.15)			(25.75)
Willkie (= 1)		−16.11		6.88			
		(27.11)		(27.71)			
Mobilization rate (%)			−60.02	−58.16			
			(287.50)	(275.10)			
Female education (years)					33.95**		40.09***
					(14.29)		(13.38)
Female employment (%)						400.93*	−1.99
						(229.63)	(325.41)
R^2	.127	.008	.001	.130	.109	.062	.277

Note.—$N = 48$. All regressions are weighted by the state's total population. The dependent variable in each regression is total state-level Lanham Act child care spending per child aged 0–12 (expressed in real 2012 dollars). The regressions omit Alaska, the District of Columbia, and Hawaii. Standard errors are in parentheses.
* $p < .10$.
** $p < .05$.
** $p < .01$.

the female employment variable becomes statistically insignificant. Nevertheless, this model explains 28% of the variation in child care spending. The coefficient on coastal location in this and the other combined regression in column 4 is statistically significant and of a similar magnitude.[11]

IV. Analysis of Maternal Employment

A. Data Description

The data set for the maternal employment analysis is crafted from the 1% IPUMS of the 1940 and 1950 US Decennial Censuses (Ruggles et al. 2015). The sample includes white and black non-Hispanic women in their prime working years (ages 25–64).[12] Sample women are retained regardless of their

[11] In results not reported here, I examine several other state characteristics, including the fraction of individuals residing on farms, the fraction black, the fraction of married women, the fraction residing in urban areas, and the fraction US-born. Although these variables are sometimes statistically significant, they do not alter the coastal state effect.

[12] Evidence in US Senate (1943a) and Fousekis (2011) indicates that a nontrivial number of black women were employed in war-related industries throughout the nation, even in the South. Furthermore, many black children attended Lanham Act child care centers. Interestingly, in a 1943 hearing of the Senate Committee on Education and Labor, it was stated that 259 "Negro"-only centers were in operation throughout the nation, in addition to many others that were integrated (US Senate 1943a). Therefore, the baseline maternal employment model is estimated on white and black women. Subgroup estimations are also conducted separately by race.

place of birth as well as the presence and number of children in the household. Not included in the analysis are women residing in group quarters or on farms. Since 1940 and 1950 IPUMS data do not include individuals in Alaska and Hawaii, women from these states are not represented in the analysis. After pooling observations for both Census dates, the analysis sample is a repeated cross section of 450,774 women (1940: 205,516; 1950: 245,258).

I examine two employment outcomes related to the extensive and intensive margins of labor supply. I begin by exploring a binary indicator of employment status that equals unity if a given woman engaged in any paid labor during the Census-defined reference week. In 1940, the reference week was March 24–30 (the Census was conducted on April 1), while the reference week in 1950 was the "previous week" (the Census was conducted over a longer time period). To be coded as employed, women must have worked at least 1 hour for pay or profit during the reference week, worked at least 15 hours as an "unpaid family worker," or had a job from which they were temporarily absent (e.g., due to illness or vacation). I then consider a measure of the total number of hours each woman was at work during the reference week. This variable has a top-code of 98 hours in both Censuses.

I identify the employment response to the Lanham Act in part by exploiting the program's age eligibility rules regarding children. Recall that the Lanham Act served women with children ages 0–12. Therefore, I define the treated group—or those likely to be exposed to the Lanham Act—as women whose youngest child was in the 0–12 age range during the period in which the Lanham Act was in operation. Such children were ages 4–19 in the 1950 Census. Two sets of women comprise the comparison group, or those not likely to be exposed to the Lanham Act: (i) women whose youngest child was age 13–17 during the program's operation, making these children age 20–24 in the 1950 Census, and (ii) childless women ages 25–64 in the 1950 Census. Women in the former group were ineligible to receive child care because their children were above the age eligibility threshold, while those in the latter group were ineligible because they did not have children. Individuals in the same age groups were identified in the 1940 Census—making this the pre-reform period—while the 1950 Census is designated as the post-reform period.

It is important to highlight a few cautions regarding the construction of the comparison groups. The Census variable used to identify the first comparison group (mothers whose youngest child is age 20–24) is based on the age of the youngest child residing within the household. This creates a potential challenge given that children likely begin leaving the household at some point within this age range. The implication is that the subset of mothers who have children ages 20–24 still residing in the home may not be representative of all mothers with children in this age range. The lack of representativeness is problematic insofar as there are changes over time in the characteristics of mothers with older children still in the household or that these characteristics are correlated with states' Lanham Act spending. In addition, there is a concern with the second comparison group (childless women). The pre-reform employment rate may be very high for this group, leaving little space available for there to be cross-state changes over time. However, the pre-reform employment rate for this group is approximately 45%, and there is substantial cross-state variation, ranging from 32% to 64%. Thus, such a "ceiling" effect is unlikely to be problematic for these women.[13] In results not reported in the paper, I estimate separate DDD models that use each comparison group individually. Estimates from these models are statistically significant and similar in magnitude.

Table 2 displays summary statistics for the employment outcomes (panel A) as well as the demographic characteristics (panel B) of women in the treated and comparison groups before and after the implementation of the Lanham Act. The employment rate for women in the treated group rose from approximately 18% in 1940 to 27% in 1950, corresponding to a 50% increase. Those in the comparison group witnessed a smaller rise in employment, from 41% to 47%, or a 15% increase. Column 5 confirms this differential employment increase through a DD-type analysis (which conditions on the demographic characteristics). The DD estimate implies that the employment rate for treated women increased 1 percentage point between 1940 and 1950. A similar story holds for the measure of hours of work, for which the DD estimate shows almost a 1-hour increase in the number of hours worked per week.

As shown in panel B, treated women are slightly younger, on average, than their counterparts in the comparison group, and they are more likely to be white, married, and to have no more than a high school degree. Columns 5 and 6 provide a series of "balance" tests on the demographic controls by estimating simple DD as well as DDD regressions. The DD estimates in column 5 suggest that there were small changes over time in the relative characteristics of women in the treated and comparison groups. Specifically, the DD estimate is small and statistically insignificant in the models for marital status, educational attainment, and between-state relocation (since childbirth), and it implies substantively minor changes in age, race, and metropolitan residence. In addition, the DDD estimates in column 6 reveal that Lanham Act spending generally had small and statistically insignificant effects on the demographic characteristics. Together, this evidence suggests that the treated and comparison groups did not experience large, differential compositional changes between the pre- and post-reform periods, thus bolstering confidence in the ability of the comparison group to provide a valid counterfactual. This is particularly comforting in the case

[13] It is also important to point out that the employment rate for childless women increased from 45% to 50% between 1940 and 1950, providing further evidence against a ceiling effect in the employment rate.

Table 2
Summary Statistics for Women in the Treated and Comparison Groups

Variable	1940 Census: Pre-reform Treated (1)	1940 Census: Pre-reform Comparison (2)	1950 Census: Post-reform Treated (3)	1950 Census: Post-reform Comparison (4)	DD Estimate (5)	DDD Estimate (6)
A. Employment outcomes:						
Employed (%)	.175	.413	.271	.466	.010***	
	[.380]	[.492]	[.445]	[.499]	(.003)	
Hours of work	6.22	16.35	9.97	18.03	.725***	
	[15.51]	[21.91]	[18.05]	[21.15]	(.126)	
B. Demographic characteristics:						
Age (years)	40.63	42.63	40.05	44.54	−1.493***	−.0038**
	[8.27]	[11.87]	[7.91]	[11.51]	(.155)	(.0019)
White (%)	.934	.896	.917	.868	.009**	.0001***
	[.249]	[.305]	[.277]	[.338]	(.004)	(.0000)
Married (%)	.897	.599	.889	.610	−.006	−.0000
	[.305]	[.490]	[.314]	[.488]	(.004)	(.0000)
High school or less (%)	.910	.862	.874	.842	.002	.0000
	[.287]	[.345]	[.332]	[.365]	(.004)	(.0001)
Foreign born (%)	.179	.130	.092	.112	−.036***	.0001
	[.384]	[.336]	[.289]	[.315]	(.004)	(.0001)
Metropolitan residence (%)	.633	.675	.620	.694	−.026***	.0002***
	[.482]	[.469]	[.486]	[.461]	(.005)	(.0001)
Ever moved states (%)	.271	.297	.303	.314	−.002	.0001
	[.444]	[.457]	[.460]	[.464]	(.005)	(.0001)

SOURCE.—Author's analysis of the 1940 and 1950 US Decennial Censuses (IPUMS).
NOTE.—Treated women are those whose youngest child is age 4–19 in the 1940 and 1950 Censuses. Comparison women are those whose youngest child is age 20–24 in the 1940 and 1950 Censuses and those who are childless. The means for weekly hours of work are not conditioned on having nonzero hours. The DD estimates in col. 5 are generated by the following equation: $Y_{ist} = \beta_1 \text{treated}_{ist} + \beta_2 \text{post}_t + \beta_3(\text{treated}_{ist} \times \text{post}_t) + Z'\psi + \mu_{ist}$. The DDD estimates in col. 6 are generated by the following equation: $Y_{ist} = \beta_1 \text{treated}_{ist} + \beta_2 \text{post}_t + \beta_3 \text{lanham}_s + \beta_4(\text{treated}_{ist} \times \text{post}_t) + \beta_5(\text{treated}_{ist} \times \text{lanham}_s) + \beta_6(\text{post}_t \times \text{lanham}_s) + \beta_7(\text{treated}_{ist} \times \text{post}_t \times \text{lanham}_s) + Z'\psi + \mu_{ist}$. All regressions control for the demographic characteristics listed in panel B. Standard deviations in cols. 1–4 are in squared brackets; standard errors in cols. 5 and 6 are in parentheses.
** $p < .05$.
*** $p < .01$.

of between-state movers, given the evidence that some families moved to war impact areas—especially those located in California—to work in war-related industries (Johnson 1993).[14] Indeed, as shown in table 2, cross-state mobility

[14] Fousekis (2011) estimates that California's population increased 30% between 1940 and 1945, with most of the new residents settling in San Diego, Los Angeles, and the Bay Area. Johnson (1993) shows that the Midwest and the South provided many of California's new residents, who apparently were responding to advertisements promising higher-paying jobs in war-related industries.

rates increased between 1940 and 1950, but they did so for all women in the sample, suggesting that war-driven mobility was a broadly shared phenomenon.

B. Empirical Specification

To estimate the impact of the Lanham Act on maternal employment, I utilize a difference-in-difference-in-differences (DDD) estimator. The model exploits three sources of variation in the data: (i) the program's age eligibility rule, which stipulated that women with children ages 0–12 could receive child care, (ii) geographic (i.e., cross-state) variation in the generosity of Lanham Act expenditures, and (iii) temporal variation in exposure, defined by distinct pre- and post-reform periods. Together, these sources of variation lead to the following empirical model:

$$\begin{aligned} Y_{ist} = {} & \beta_1 \text{post}_t + \beta_2(\text{treated}_{ist} \times \text{post}_t) + \beta_3(\text{treated}_{ist} \times \text{lanham}_s) \\ & + \beta_4(\text{post}_t \times \text{lanham}_s) + \beta_5(\text{treated}_{ist} \times \text{post}_t \times \text{lanham}_s) \\ & + \boldsymbol{Z}'\psi + \boldsymbol{S}'\gamma + \upsilon_s + \mu_{ist}, \end{aligned} \tag{1}$$

where Y is an employment outcome for woman i in state s and year t. As previously stated, the employment outcome refers to the binary indicator of employment status or the measure of weekly hours-of-work. The variable post is a binary indicator that equals unity for observations drawn from the 1950 Census (and zero for observations in the 1940 Census), while the variable treated is a binary indicator that equals unity if the ith woman's youngest child was aged 0–12 during the Lanham Act's operation (and zero for women in the comparison group). The variable lanham is the measure of state-level Lanham Act spending per child aged 0–12. The vector given by $\boldsymbol{Z}'$ is a set of demographic controls, including age (fixed effects), race (one dummy variable), marital status (six dummy variables), educational attainment (21 dummy variables), a binary indicator for foreign born, and metropolitan residence (four dummy variables). Importantly, the model also controls for cross-state relocation through a binary indicator that equals unity if a given women's current state of residence differs from her state of birth.[15] Also included in equation (1) is a set of state-of-residence fixed effects (υ_s)

[15] In robustness checks, I estimate two additional models that attempt to control for cross-state mobility. First, I estimate the DDD model on the subset of women whose current state of residence is the same as the state of birth (i.e., adult women who reside in the same state in which they were born). Second, following Acemoglu et al.'s (2004) approach to account for mobility, I estimate a DDD model that includes state-of-birth fixed effects as well as current-state-of-residence fixed effects. Estimates from these specifications are quite similar to those reported in the paper.

to control for permanent differences across states that may be correlated with Lanham Act spending. Equation (1) is estimated using ordinary least squares (OLS) regression, and the standard errors are adjusted for arbitrary forms of heteroskedasticity as well as state-level clustering.[16]

The coefficient of interest is β_5, which provides the DDD estimate of the effect on maternal work activity of an increase in Lanham Act spending for treated relative to comparison group women, after versus before reform. It is important to be clear about what β_5 captures. Given that the post-reform outcomes are observed in 1950—approximately 3 years after the Lanham Act was terminated—β_5 reflects the medium-run employment response to the child care program. In addition, β_5 is interpreted as the intent-to-treat (ITT) effect of the Lanham Act. As such, it captures the reform effect on at least two groups of women with age-eligible children. First, there is the direct effect of the child care reform on mothers who used Lanham Act centers during the war. For some of these women, the short-run increase in employment translated into a longer-run response. Second, there is the indirect effect on women who utilized forms of non–Lanham Act child care during the war or who began using child care after the war.[17] As previously stated, the Lanham Act altered society's views regarding institutional child care; this cultural shift likely catalyzed some women to combine work with other forms of child care during the war as well as after the program was terminated.

As with all DD designs, the model specified in equation (1) must overcome two key threats to the unbiasedness of β_5. The first threat is the presence of unobserved, contemporaneous shocks that may have differentially affected maternal employment across states spending different amounts on child care. The obvious concern in this context is the shock represented by the war, which by itself produced a sizable employment response among women (Acemoglu et al. 2004; Goldin and Olivetti 2013). I do several things

[16] In results not reported in the paper, I estimate a DDD model that includes a control for the number of children residing in the household. The results are similar to those in the baseline model. In addition, I estimate the model with state-year clustered standard errors. The standard errors are consistent with those in the baseline model.

[17] There is scattered evidence from a variety of sources that many mothers used non–Lanham Act child care during the war. The principal reason for this is that communities generally did not receive enough funding to serve all children in need of services. As described in Fousekis (2011), it was common for communities to receive far less in grants from the federal government than the amount applied for. Evidence from local needs assessments conducted during the war indicate that mothers relied on a patchwork of neighbors and relatives to provide child care in the absence of Lanham Act care. There is also evidence that existing child care centers (which traditionally served low-income families and which did not receive Lanham Act funding) expanded their services, to the fullest extent possible, to accommodate the children of mothers working in wartime industries (Fousekis 2011).

to disentangle the war's effect on employment from that of the Lanham Act. My primary strategy, as shown in equation (1), is to interact the variables post and lanham with the variable treated; in doing so I compare the effect of Lanham Act spending across two groups women (i.e., treated and comparison group women) whose labor supply was arguably influenced in the same manner by the war but only one of which (the treated group) was influenced by the child care program. Another strategy is to control directly for any war-driven changes in employment. I do so by including in S' the state-level Armed Forces mobilization rate (interacted with variables treated and post). As previously discussed, this variable proxies the outflow of men from the labor market and the rising need for female workers.

The second threat to the validity of the DDD estimate is the possibility that states with different pre-existing trends in maternal employment pursued Lanham Act funding more or less aggressively.[18] Alternatively, the federal government might have made child care funding decisions by considering long-standing differences in the stock of human capital between states. Indeed, the regression results presented in table 1 indicate that the federal government may have favored states with higher levels of educational attainment. Once again, I use the interaction of post and lanham with treated as the primary tool for handling the nonrandom assignment of child care funds to states. In particular, the three-way interaction should purge the effect of any differential employment trends (across states that received different levels of funding) because childless women and those with older children were not eligible to receive child care funds. I also incorporate in the baseline model a number of state-level characteristics, S', including a dummy variable for women's location in a coastal state, the female employment rate, average female educational attainment (in years), the fraction black, the fraction of individuals residing on farms, the fraction of women who are married, the fraction residing in urban areas, the fraction US born (all interacted with treated and post). These variables are estimated from the 1940 Census, and they should help to control for pre-existing trends in employment as well as cross-state differences in Lanham Act spending.

[18] A related concern is whether the DDD model can distinguish Lanham Act–driven increases in employment from the more general rise in employment that may have occurred in the absence of the Lanham Act. Specifically, it is possible that the DDD coefficient is identified off changes over time in the interaction between treated and lanham. To explore this possibility, I estimate a DD model of the employment outcomes on treated, lanham, treated × lanham, and the full set of demographic controls using only the 1940 Census. In the models of any work and hours of work, the DD coefficient is small in magnitude and statistically insignificant, suggesting that the DDD estimates reported in the paper are not picking up the effect of general changes in employment but instead reflect the effect of the Lanham Act's implementation.

C. DDD Estimation Results

Table 3 presents results from the DDD model. Columns 1 and 2 provide estimates for the binary indicator of employment, and columns 3 and 4 provide those for weekly hours-of-work. The odd-numbered columns omit the state-level controls, while the even-numbered columns include them. Therefore, columns 2 and 4 should be considered the baseline specification.

The DDD estimates show a positive effect of the Lanham Act at both work margins. Comparing columns 1 and 2 and columns 3 and 4, it appears that adding the state-level controls does not influence the results. In terms of the extensive margin, the baseline estimate implies that a $1 increase in Lanham Act spending led to a 0.1 percentage point increase in the employment rate. One way to interpret this estimate is to calculate the change in employment when the average treated woman is moved from a low- to a high-spending state. For example, states in the bottom quartile of the distribution spent approximately $22 per child, while those in the top quartile spent about $66 per child. Applying these figures to the DDD estimate implies that a move from the bottom to the top quartile of Lanham Act spending would generate a 4.4 percentage point increase in the employment rate for treated women. This estimate is one-quarter of the pre-reform mean for the treated group. A similar story emerges in the models examining labor supply at the intensive margin. The baseline DDD estimate suggests that a $1 increase in Lanham Act spending increased weekly hours of work by 0.05 hours. In other words, moving the average treated woman from the bottom to the top quartile of Lanham Act spending would generate an additional 2 hours of work per week. This effect is about one-third of the pre-reform mean for the treated group.

Although it is not reported in table 3, the DD coefficient on $\text{post}_t \times \text{lanham}_s$, β_4, is also noteworthy. Specifically, it assesses the impact of the Lanham Act on the employment change experienced by women in the comparison group. It suggests that employment—as measured by any work and weekly hours-of-work—declined by a small amount following the enact-

Table 3
Main Difference-in-Difference-in-Differences (DDD) Employment Results

Variable	(1)	(2)	(3)	(4)
Treated × post × lanham	.0010***	.0010***	.0457***	.0450***
	(.0003)	(.0003)	(.0143)	(.0133)
Outcome	Employed	Employed	Hours of work	Hours of work
State controls?	No	Yes	No	Yes
Outcome mean	.175	.175	6.22	6.22

SOURCE.—Author's analysis of the 1940 and 1950 US Decennial Censuses (IPUMS).
NOTE.—Observations = 450,774. Standard errors (in parentheses) are adjusted for state-level clustering. Columns 1 and 3 do not include the state-level controls, while cols. 2 and 4 add the state-level controls discussed in the text. The outcome mean refers to the pre-reform employment mean for the treated group.
*** $p < .01$.

ment of universal child care.[19] This indicates the possibility that women with young children began to crowd out their counterparts with older children and those without children. It is also possible that, once a guarantee of child care was established, women with young children were more attractive to employers than other women in the labor market. In any case, this result suggests that increases in employment for one group of women may have come at a cost of reduced employment for other groups.

Table 4 examines a number of subgroups, as defined by the age of the youngest child and the race, age, marital status, and educational attainment of the woman. I also test the null hypothesis of equal DDD coefficients across subgroups within each characteristic. As shown in the first two rows, treated women with younger children (ages 0–5) were more responsive to the Lanham Act than those with older children (ages 6–12). This pattern is consistent with the contemporary literature estimating the labor supply effects of child care prices and subsidies (Anderson and Levine 2000; Baker et al. 2008). Given these results, it is not surprising that younger women (ages 25–44) were more responsive to the Lanham Act than older women (ages 45–64). Acemoglu et al.'s (2004) work similarly finds that war mobilization had a larger impact on the employment of younger women. It also appears that the Lanham Act increased the employment of white and African American women, a finding that accords with Fousekis's (2011) discussion that many black children were served in the war-time child care centers. The remaining results in table 4 reveal smaller differences by marital status and educational attainment. That lower-and higher-skilled women responded nearly identically to the Lanham Act is intriguing in light of results in Acemoglu et al. (2004) and Goldin and Olivetti (2013). Both papers find that war mobilization effects were concentrated on higher-skilled women while having virtually no effect on the less skilled.

V. Life-Cycle Analysis of Long-Run Outcomes

A. Data Description

To examine the long-run impact of the Lanham Act, I use 1970 (1%), 1980 (5%), and 1990 (5%) Decennial Census data from the IPUMS (Ruggles 2015). I constrain the analysis to white and black non-Hispanic individuals who were born within the continental United States and who were not residing in group quarters. I drop individuals for whom information on state of birth is not available, since Lanham Act spending data and other state-level variables cannot be merged to them. As will be explained in more detail below, I restrict the sample to individuals meeting year-of-birth cri-

[19] In the model of any work, the coefficient (standard error) on $\text{post}_t \times \text{lanham}_s$ is $-.0004$ (.0002); in the model of weekly hours-of-work, the coefficient (standard error) is $-.0213$ (.0076).

Table 4
Difference-in-Difference-in-Differences (DDD) Subgroup Analyses

Variable	Coefficient (1A)	SE (1B)	Coefficient (2A)	SE (2B)
Youngest child: ages 0–5	.0013***	.0004	.0583***	.0151
Youngest child: ages 6–12	.0006**^^^	.0002	.0300***^^^	.0107
White women	.0009***	.0003	.0423***	.0114
Black women	.0007**	.0004	.0513***	.0166
Women ages 25–44	.0012***	.0004	.0574***	.0191
Women ages 45–64	.0006***^^	.0002	.0317***^	.0067
Married	.0010***	.0003	.0418***	.0137
Unmarried	.0004**^^	.0002	.0360***	.0114
< High school	.0010***	.0003	.0467***	.0133
≥ High school	.0010***	.0003	.0438***	.0119
Outcome	Employed		Hours of work	

SOURCE.—Author's analysis of the 1940 and 1950 US Decennial Censuses (IPUMS).
NOTE.—Each cell shows the coefficient and standard error (in parentheses and adjusted for state-level clustering) on treated × post × lanham. Columns 1A and 1B show the coefficient and standard error for the model of any employment. Columns 2A and 2B show the coefficient and standard error for the model of weekly hours of work. All models include the full set of controls outlined in the baseline model.
^^^ The given subgroup coefficient is statistically significantly different from its counterpart at the .01 level.
^^ The given subgroup coefficient is statistically significantly different from its counterpart at the .05 level.
^ The given subgroup coefficient is statistically significantly different from its counterpart at the .10 level.
** $p < .05$.
*** $p < .01$.

teria; this ensures that adults assigned to the treated group were (as children) age-eligible for child care when the Lanham Act was administered. The analyses are based on 456,070 observations in the 1970 Census, 2,500,553 observations in the 1980 Census, and 2,481,049 observations in the 1990 Census.

As shown in panel A of table 5, I examine five outcomes related to adults' labor market activity and program participation. Specifically, I examine a binary indicator of employment status that equals unity if a given individual engaged in any paid labor during the reference week. I also create an indicator of full-time employment, defined as working at least 35 hours during the reference week. The Census's retrospective questions on labor market activity and benefit receipt provide the foundation for the last three outcomes: a binary indicator for any paid labor during the previous calendar year, total pre-tax wage and salary income in the previous year, and a binary indicator for whether an individual received public assistance in the previous year. Public assistance is the sum of income from Supplemental Security Income (SSI), Aid to Families with Dependent Children (AFDC), and General Assistance (GA).

Given the large number of outcomes, all of which are observed at each Census date (for a total of 15 outcomes), I follow Deming (2009) and construct a summary index of each set of five outcomes in the 1970, 1980, and

Table 5
Summary Statistics for the Adult Outcomes and Select Demographic Characteristics

Variable	1970 Census (1)	1980 Census (2)	1990 Census (3)
A. Labor market and program participation outcomes:			
Employed (%)	.666	.756	.772
	(.472)	(.429)	(.420)
Employed full-time (%)	.824	.851	.851
	(.381)	(.357)	(.356)
Employed last year (%)	.774	.814	.826
	(.418)	(.389)	(.379)
Earnings last year (2014 dollars)	36,183	47,458	53,564
	(28,789)	(36,103)	(50,358)
Public assistance (%)	.023	.033	.028
	(.150)	(.180)	(.164)
B. Demographic characteristics:			
Age (years)	28.31	37.84	47.82
	(6.03)	(6.10)	(6.10)
Male (%)	.475	.489	.486
	(.499)	(.500)	(.500)
White (%)	.890	.886	.905
	(.313)	(.318)	(.294)
Married (%)	.726	.772	.756
	(.446)	(.419)	(.430)
Ever moved states (%)	.464	.389	.396
	(.499)	(.487)	(.489)

Source.—Author's analysis of the 1970, 1980, and 1990 US Decennial Censuses (IPUMS).
Note.—Standard deviations are in parentheses.

1990 Censuses (for a total of three outcomes). I normalize the individual outcomes to have a mean of zero and a standard deviation of one. After reversing the signs on a few—so that positive values indicate an improvement in well-being—I create a new variable representing the mean score of the normalized outcomes. Thus, higher values on the summary index imply better labor market and program participation outcomes. This approach is advantageous because the effect sizes estimated here can be compared with those from Deming's (2009) analysis of Head Start and Anderson's (2008) reevaluation of the Perry and Abecedarian programs. Nevertheless, I present results from the summary index alongside the estimates from the five individual outcomes at each Census date.

To identify the impact of the Lanham Act on adult outcomes, I estimate a DD model similar in spirit to Havnes and Mostad (2011b). In particular, I draw comparisons of the adult outcome indices between cohorts exposed and not exposed to the Lanham Act as children, once again taking advantage of the large cross-state variation in the generosity of Lanham Act expenditures. Recall that the child care program operated during the period 1943–

46, with eligibility extending to children ages 0–12. Therefore, treated cohorts—defined as individuals who were age-eligible for the child care program for at least 1 year—were born between 1931 and 1946. These cohorts were ages 24–39 in the 1970 Census, ages 34–49 in the 1980 Census, and ages 44–59 in the 1990 Census. Thus, by measuring adult outcomes for these cohorts at three Census dates, this study is able to trace the full life-cycle effects of the Lanham Act. To construct a comparison group of individuals not exposed to the Lanham Act, I include cohorts born in the first 5 years after the Lanham Act was terminated. Specifically, these individuals were born during the period 1947–51, making them ages 19–23 in the 1970 Census, ages 29–33 in the 1980 Census, and ages 39–43 in the 1990 Census.[20] Therefore, the analysis sample is restricted to include individuals ages 19–39 in the 1970 Census, ages 29–49 in the 1980 Census, and ages 39–59 in the 1990 Census.

One may inquire about the choice of using cohorts born after the Lanham Act's termination as the comparison group, as opposed to cohorts born before the Lanham Act was initiated. The key disadvantage of using pre-Lanham Act cohorts is that a sizable number of depression-era nursery schools were operated by the federal government in the years preceding the war. Although the child care centers were defunded after the war began, many of them eventually reopened in order to serve the goals of the Lanham Act. Thus, children in the pre–Lanham Act cohorts were potentially exposed to a system of federal child care that, in some ways, was quite similar to the system operating under the Lanham Act. If the counterfactual in-

[20] Given that individuals in the treated and comparison cohorts are quite close in age, one concern is that some in the comparison cohorts were partially, though indirectly, treated by the Lanham Act because they had one or more older siblings in the treated cohorts. In other word, while those in the comparison cohorts did not attend a Lanham Act child care center, the adult outcomes of such children might have been influenced through changes in maternal employment, family income, or some other channel. To the extent that comparison cohorts were treated, it means that the DD estimates on the Lanham Act are likely to be understated. I do a few things to investigate the extent and implications of this issue. First, I use the 1950 Census to calculate the share of children born between 1947 and 1950 (i.e., in a comparison cohort) who have at least one sibling born between 1931 and 1946 (i.e., in a treated cohort). This figure is 58%. Although this a nationwide figure based on the full 1950 Census (as opposed to the analysis sample described in the previous section), it does suggest that a nontrivial number of those in the comparison cohorts might have been indirectly treated by the Lanham Act. Second, I use the 1970 Census to estimate the DD model using the subset of comparison cohorts without a sibling present in the household. Although removing only those with adult co-resident siblings is imperfect, it is not possible in the Census to identify all individuals with siblings. I use the 1970 Census to ensure the closest correspondence between those in the comparison cohorts with co-resident siblings and those with any siblings. The DD estimate in this alternative model is quite close to that in the baseline model discussed in the text.

cludes children who also attended child care, then using the pre–Lanham Act cohorts may yield DD estimates that understate the impact of the Lanham Act.

Although the use of post–Lanham Act cohorts surmounts this particular issue, there is a potential drawback as well from using these individuals as the comparison group. Given that the post–Lanham Act cohorts are younger than those in the treated cohorts, it is possible that, in households with two or more children, the outcomes of younger (comparison group) siblings were influenced by spillovers from changes in maternal time and material inputs as well as through peer effects from their older (and treated) siblings. However, as with the pre–Lanham Act comparison cohorts, the net effect of such spillovers would be to yield DD estimates that understate the impact of the Lanham Act. Although the primary analysis uses post–Lanham Act cohorts as the comparison group, I implement a robustness check that tests an alternative comparison group comprised of cohorts born far enough before the Lanham Act was enacted that they were never age-eligible to receive child care. Results from this DD analysis are comparable to those from the main specification using the post–Lanham Act cohorts.

B. Empirical Specification

The DD model exploits two sources of variation: (i) the Lanham Act's age eligibility rule, which, in conjunction with individuals' year of birth, creates the treated and comparison cohorts described above, and (ii) geographic (i.e., cross-state) variation in the generosity of Lanham Act expenditures. These sources of variation are used to estimate the following regression model:

$$Y_{its} = \beta_1 \text{treated}_{ist} + \beta_2(\text{treated}_{ist} \times \text{lanham}_s) + \boldsymbol{Z}'\psi + \boldsymbol{S}'\gamma + \upsilon_s + \mu_{ist}, \quad (2)$$

where Y is the outcome summary index for person i in cohort t and state-of-birth s. The variable treated is a binary indicator that equals unity if a given adult was exposed to the Lanham Act as a child (i.e., born between 1931 and 1946), and zero if the adult was not exposed to the program (i.e., born between 1947 and 1951). The variable lanham is the measure of state-level Lanham Act spending. It is merged to individuals based on the state-of-birth reported in the Census. The vector given by $\boldsymbol{Z}'$ includes a number of demographic controls, such as gender, race (one dummy variable), marital status (five dummy variables), farm residence, and region of residence (14 dummy variables). Also included is a proxy for individual's mobility, defined as a binary indicator for whether the state of residence (at the time the Census was taken) and the state of birth are the same. The model is estimated separately on the 1970, 1980, and 1990 Censuses. Standard errors are adjusted for arbitrary forms of heteoskedasticity as well as state-level clustering.

The impact of the Lanham Act is given by the coefficient β_2, which is interpreted as the effect of increased child care spending for adults in treated cohorts relative to adults in the comparison cohorts. This DD estimate is conceptually equivalent to a comparison of the difference in outcomes between treated and comparison cohorts born in high-spending states with the difference for these groups born in low-spending states. It should be interpreted as the ITT effect of the Lanham Act. That is, this parameter averages the impact of reform over treated adults who received and did not receive Lanham Act child care. The ITT is an interesting parameter in this context for several reasons. First, it is useful for characterizing the impact of policy reforms that do not serve the full eligible population. That the Lanham Act did not reach full penetration makes the DD estimates in this analysis more relevant to the current early care and education landscape, in which most programs also do not fully serve the eligible population. Second, the ITT captures the reform effect on exposed nonparticipants in a program (i.e., noncompliers) who nevertheless use similar services in part because of the reform in question. This is relevant for policies like the Lanham Act, which apparently catalyzed a deep cultural shift in the way Americans viewed institutional child care. This shift may have spurred some families that did not receive Lanham Act child care to use other forms of nonparental arrangements during and after the war. The DD estimates additionally capture the long-run implications of utilizing these non–Lanham Act child care arrangements.[21]

Given that the baseline DD model includes a set of state-of-birth fixed effects (υ_s), it is unnecessary that Lanham Act spending be orthogonal to the demographic and economic characteristics of states. Rather, to be interpreted as causal, the estimation of β_2 must overcome the two identification challenges. First, there could be unobserved contemporaneous shocks (e.g., the war) that may differentially affect adult outcomes across states spending different amounts through the Lanham Act. The primary strategy for purging β_2 of any war-induced changes in adult outcomes is to include in $\boldsymbol{S}'$ the state-level Armed Forces mobilization rate. Also included in $\boldsymbol{S}'$ are a number of observable state characteristics that may differentially affect outcomes in high- and low-spending states. These variables include a binary indicator for children's birth in a coastal state, the female employment rate, the fraction black, the fraction of individuals residing on farms, the fraction of women who are married, the fraction residing in urban areas, and the fraction US-born. Finally, I control for state-level Lanham Act spending on schools and recreation to account for the fact that the legislation funded other public works projects that may influence short- and long-

[21] Please refer to footnote 18 for a full discussion of the use of non–Lanham Act child care arrangements during the war.

run outcomes.[22] All of the state-level controls are interacted with the variable treated.

The second challenge is the possibility that adult outcomes for individuals born in high- and low-spending states would have trended differently in the absence of the Lanham Act. To examine whether the common trends assumption holds, I exploit the fact that the comparison cohorts are drawn from a (post-reform) period in which the Lanham Act no longer operated. Given that the Lanham Act was terminated in 1946, I use the comparison cohorts—who were born between 1947 and 1951—to explore the outcome trends for those born in states that spent different amounts on child care. Figure 2, parts *A*, *B*, and *C*, graph the mean on the outcome summary index by post–Lanham Act birth cohort separately for states in the first through the fourth quartile of the Lanham Act spending distribution. Each figure corresponds to the 1970, 1980, and 1990 Censuses, respectively. The figures reveal a close correspondence in the comparison group outcome trends throughout the distribution of child care spending.

Despite this evidence, I conduct several auxiliary analyses to ensure that the DD framework satisfies the common trends assumption. First, I conduct a matched case control analysis, in which each high-spending state (i.e., those in the top quartile of the Lanham Act spending distribution) is paired with one or more low-spending states (i.e., those in the bottom quartile of the spending distribution) based on the degree of similarity in the outcome trends. Once a set of comparison states is identified, I estimate a DD model that examines the difference in adult outcomes between treated and comparison cohorts in states where spending was high and similarly trending states where spending was low. Second, I implement a placebo reform that "turns on" the Lanham Act for some cohorts in the comparison group. If the adult outcomes follow a similar time path in low- and high-spending states, then the estimated effect of the placebo reform should be statistically insignificant. Third, I estimate a DD model that includes state-specific linear time (i.e., cohort) trends. This analysis allows for differential outcome trends for those in the treated and comparison cohorts residing in states that spent different amounts on child care. Finally, I estimate the DD model using an alternative comparison group: cohorts born prior to the Lanham Act's implementation. Recall that the baseline DD model draws its comparison group from cohorts born after the Lanham Act was terminated. In contrast, the alternative comparison group is comprised of individuals born far enough before the Lanham Act that they were never age-eligible to receive child care.

[22] It is important to note that the expenditure data for these categories are less complete than those for the child care program; only data for FY1945 are available. Average per child expenditures on schools was $103 and that on recreation facilities was $9 in FY1945.

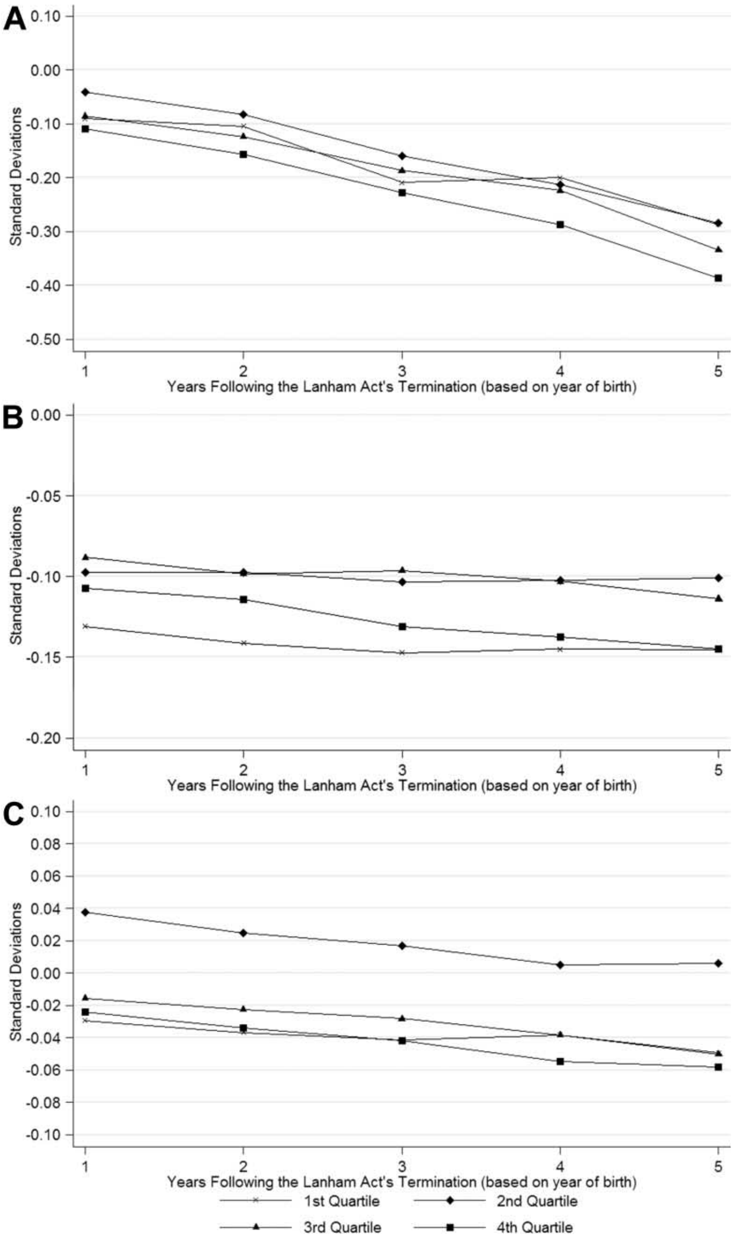

FIG. 2.—Mean score on the outcome index, comparison cohorts by Lanham Act spending quartile: *A*, 1970 US Decennial Census; *B*, 1980 US Decennial Census; *C*, 1990 US Decennial Census.

C. Baseline DD Estimates

Results from the DD analysis are shown in table 6. Columns 1–3 show the DD estimates based on the 1970, 1980, and 1990 Censuses, respectively, along with the standard errors (in parentheses). Panel A presents the DD estimates based on the outcome summary index, while panel B provides the results for the individual labor market and program participation outcomes.

Looking first at the models in the top row of panel A, which includes only the state-of-birth fixed effects and the state-level controls, the estimates reveal that the Lanham Act had beneficial and persistent long-run effects. Indeed, the DD coefficients are consistently positive and highly statistically significant, indicating that increases in Lanham Act spending led to better adult outcomes for the treated cohorts relative to the comparison cohorts. The DD estimates also reveal fairly steady treatment effect sizes that persisted throughout adulthood. Adding the demographic covariates does little to alter the DD estimates, as shown in the second row of panel A. Thus, results in the second row are considered the baseline DD estimates.

Table 6
Baseline Difference-in-Differences (DD) Results for the Impact of the Lanham Act

	Census Sample (1970) (1)	Census Sample (1980) (2)	Census Sample (1990) (3)
A. Outcome summary index:			
State-level controls only	.00033***	.00016***	.00020***
	(.00007)	(.00003)	(.00004)
All controls	.00035***	.00012***	.00017***
	(.00007)	(.00003)	(.00004)
B. Individual outcomes:			
Employed	.00012***	.00003*	.00007**
	(.00004)	(.00002)	(.00003)
Employed full-time	.00041***	.00004**	.00002
	(.00005)	(.00002)	(.00001)
Employed last year	.00008**	2.80e-06	.00005**
	(.00004)	(.00002)	(.00002)
Ln(earnings last year)	.00062***	.00019***	.00025***
	(.00019)	(.00007)	(.00005)
Public assistance	−.00003***	−.00004***	−.00002***
	(.00000)	(.00000)	(.00000)
Treated cohorts' age range	24–39	34–49	44–59

Source.—Author's analysis of the 1970, 1980, and 1990 US Decennial Censuses (IPUMS).
Note.—Displayed in each cell is the coefficient on treated × lanham and its standard error (in parentheses), which is adjusted for state-level clustering. The models in the first row of panel A include only the state-level controls (as well as the state of birth fixed effects), while those in the second row add the individual-level demographic controls. The models in panel B include the full set of state-and individual-level controls.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

The DD estimate in the 1970 Census, which includes treated adults in their mid-20s to late-30s, implies that a $1 increase Lanham Act spending (per child ages 0–12) led to a 0.00035 SD improvement in the outcome summary index. The corresponding effect sizes from the 1980 Census (i.e., treated adults in their mid-30s to late-40s) and 1990 Census (i.e., treated adults in their mid-40s to late-50s) are 0.00012 SDs and 0.00017 SDs, respectively. To put these results into perspective, I use the DD estimates to simulate the effect of increasing Lanham Act spending from the bottom to the top quartile of the distribution (i.e., from $22 to $66 per child). Such an increase in child care spending would generate improvements in the outcome summary index of 0.015 SDs, 0.005 SDs, and 0.008 SDs, respectively, in the 1970, 1980, and 1990 Censuses. Taking the weighted average (according to the sample size) of all three produces an overall effect size of approximately 0.007 SDs. Another approach is to simulate the impact of a spending increase from the Lanham Act median ($44 per child ages 0–12) to the level of Georgia's pre-K program ($2,066 per 4-year-old).[23] Georgia's program is chosen because its resource ranking is near the middle of the state distribution (NIEER 2012). Such an increase in child care spending translates to an increase in the outcome summary index of 0.33 SDs (weighted average).

Is it useful to compare this long-run effect size to those generated by other early childhood programs. Deming (2009) creates a similar index of young-adult outcomes in a study of the long-run effects of Head Start. His primary finding is that Head Start participants scored 0.23 SDs higher on the outcome index relative to their siblings who did not participate. A recent paper by Anderson (2008) reanalyzes data from the Perry and Abecedarian interventions, also using a summary index of adult outcomes. The estimated effect sizes for the Perry program are 0.35 SDs for females and −0.01 SDs for males, while the corresponding effect sizes for Abecedarian are 0.45 SDs and 0.31 SDs, respectively. Thus, the effect sizes generated by the Lanham Act are comparable in magnitude to several prominent early childhood education interventions.

Turning to the individual outcomes in panel B, the results consistently point to improvements in adult labor market functioning as well as reductions in public assistance receipt. It is also noteworthy that the magnitudes of the DD estimates remain relatively constant over time. Looking at the results from the 1990 Census, I find that a $100 increase in Lanham Act spending increased the fraction currently employed by 0.7 percentage points and increased the rate of full-time employment by 0.2 percentage points. This spending increase also boosted annual earnings by 2.5%, and

[23] In 2012, Georgia spent $289,222,657 on its pre-K program (NIEER 2012). Given that the state's population of 4-year-olds totaled 140,012 in 2012, this translates to $2,066 of pre-K spending per child. The corresponding Lanham Act figure ($44 per child) is expressed in constant 2012 dollars.

it lowered the proportion of adults receiving public assistance by 0.2 percentage points.

D. Robustness Checks

Recall that a key assumption in the DD framework is that the adult outcomes for individuals born in high- and low-spending states would have the same trends in the absence of the Lanham Act. Although the analysis of comparison cohorts revealed similar time paths across the distribution of Lanham Act spending, tables A1, A2, and A3 report several additional analyses to ensure the common trends assumption is satisfied.

First, I conduct a matched case control analysis, pairing one or more comparison states with a treated state based on the degree of similarity in the outcome trends (Baum and Ruhm 2014). I begin by creating a set of treated—or high-spending—states defined as those in the top quartile of the Lanham Act spending distribution, and a set of potential comparison—or low-spending—states defined as those in the bottom quartile of the distribution. For each treated-comparison state combination, I estimate the following regression using only those in the comparison cohorts:

$$\begin{aligned} Y_{its} = {} & \beta_1 \text{trend}_{its} + \beta_2 \text{low\_spending}_{its} \\ & + \beta_3(\text{trend}_{its} \times \text{low\_spending}_{its}) + \mathbf{Z}'\psi + \varepsilon_{its}, \end{aligned} \tag{3}$$

where Y is defined in the same manner as in equation (2); trend counts from one to five for comparison group cohorts born between 1947 and 1951, and low_spending is a binary indicator that equals unity for individuals born in a potential comparison (or low-spending) state and zero for those born in a given treated (or high-spending) state. The β_3 provides an estimate of the degree to which the outcome trends in the absence of reform differ between pairs of low- and high-spending states. I include a low-spending state in the comparison group for the case control analysis if the t-statistic on β_3 does not exceed $|1.0|$. Although this cut-off is subjective, it is also quite restrictive in the sense that I ultimately discard many low-spending states whose outcome trend is not statistically different from that in a given treated state.

Once one or more comparison states are identified for each treated state, I estimate the following DD regression:

$$\begin{aligned} Y_{its} = {} & \beta_1 \text{high\_spending}_{its} + \beta_2 \text{treated}_{its} \\ & + \beta_3(\text{high\_spending}_{its} \times \text{treated}_{its}) + \mathbf{Z}'\psi + \varepsilon_{ist}, \end{aligned} \tag{4}$$

where Y is defined in the same manner as in (2), high_spending is a binary indicator that equals unity for individuals born in a given treated state and zero for those in one or more of the matched comparison states, and treated is a binary indicator that equals unity if a given adult was age-eligible for the Lanham Act as a child and zero if the adult was not age-eligible for the program. The β_3 provides the DD estimate of the difference in adult outcomes between treated and comparison cohorts born in states where Lanham Act spending was high and where spending was low.

Table A1 presents results from the DD matched case control analysis using the 1970 Census. Each row represents a separate analysis, with the treated states arrayed in descending order of Lanham Act spending (col. 1). Column 2 shows the comparison states included in each analysis, column 3 shows the number of observations, and column 4 provides the DD estimate. For example, results in the first row use California as the treated state and Arkansas, Idaho, and New Jersey as the matched comparison states. The DD estimate implies that the Lanham Act improved adult outcomes: treated cohorts in California score about 0.06 SDs higher on the outcome summary index. Positive and statistically significant effects are found in 10 of the 12 treated states, with the estimated effect sizes ranging between 0.03 and 0.10 SDs. It is noteworthy that the Lanham Act had similar-sized positive effects on individuals in geographically and economically disparate states.

In the second analysis, I implement a placebo reform that "turns on" the Lanham Act for some cohorts in the comparison group. If the adult outcomes follow a similar time path in low- and high-spending states, then the estimated effect of the placebo reform should be statistically insignificant. I conduct the placebo test by estimating the following regression:

$$\begin{aligned} Y_{\mathrm{its}} = {} & \beta_1 \mathrm{treated}_{its} + \beta_2(\mathrm{treated}_{its} \times \mathrm{lanham}_s) + \beta_3 \mathrm{placebo}_{its} \\ & + \beta_4(\mathrm{placebo}_{its} \times \mathrm{lanham}_s) + \boldsymbol{Z}'\psi + \boldsymbol{S}'\gamma + v_s + \varepsilon_{ist}, \end{aligned} \tag{5}$$

where Y, treated, and lanham are defined in the same manner as in (2). The variable placebo is a binary indicator that equals unity for individuals born in 1949 and 1950. Such individuals were incorporated in the comparison group in the main DD analysis. However, in (5), they provide the basis for the falsification test, in which placebo is interacted with lanham and β_4 provides an estimate of the differential response to increased Lanham Act spending between cohorts born in 1949 or 1950 and the cohort born in 1951. Thus, individuals born in 1951 comprise the comparison group in the placebo analysis, just as they did in the main analysis. The β_2 should be interpreted as a modified DD effect in that the treated cohorts are compared to a subset of the full comparison group.

As shown in table A2, the placebo DD estimates are never statistically significant and are substantially smaller in magnitude than their modified DD counterparts. Indeed, tests of the null hypothesis of no difference between the placebo and modified DD estimates are consistently rejected. It is noteworthy that the placebo estimates are also quite different from the baseline DD estimates in table 6, while the modified DD estimates are quite similar.

That Lanham Act spending influences those in the treated cohorts, while having no effect on individuals in the comparison group, provides additional evidence in support of the common trends assumption.

Finally, the DD estimates in table A3 come from two additional specification checks. The first row adds a set of state-specific linear time (i.e., cohort) trends. The inclusion of cohort trends explicitly allows the outcomes for treated and comparison cohorts to follow different time paths in states that spent different amounts on child care. The DD estimates remain not only statistically significant but of a magnitude similar to the baseline estimates. Recall that the baseline DD model uses post–Lanham Act birth cohorts as the comparison group. A drawback to this approach is that such children might have been treated indirectly via intrahousehold spillovers and peer effects from their older (treated) siblings. Thus, I estimate a DD model using an alternative comparison group: cohorts born far enough before the Lanham Act's enactment that they were never age-eligible to receive child care. I select cohorts born between 1928 and 1930, making them ages 13–15 in 1943, the first year in which child care was funded under the Lanham Act. As previously stated, such children were also likely to be treated, in this case by the federally administered system of child care operating during the depression. That some individuals in the pre– and post–Lanham Act comparison groups were directly or indirectly treated means that the DD estimates in both models are likely to be understated. Fortunately, results based on the alternative comparison cohorts are quite similar to those from the baseline model, as shown in the second row of table A3.

E. Distributional Effects of the Lanham Act

A central issue in the debate over the design of early care and education systems is whether a given program should be targeted at specific subgroups or universally accessible. Implicit in this discussion is the question of whether the impact of early childhood programs is experienced equally by diverse groups of children. To this point, the DD effects have been averaged over adults from low and high socioeconomic strata and with low- and high-ability endowments. Such average effects could mask substantial heterogeneity if children from different family backgrounds respond differently to early childhood experiences. Indeed, previous work finds that economically disadvantaged children generally capture moderate to large benefits from early childhood programs, while outcomes for their more advantaged peers are either not affected or even adversely affected (e.g., Deming 2009; Havnes and Mogstad 2011b; Cascio and Schanzenbach 2013). Thus, it is important for the current study to explore heterogeneity in the long-run impact of the Lanham Act.

This task is made somewhat difficult using the Census because data on (adult) respondents' early home environment or parental characteristics are not collected. The approach I take here is to examine whether the

Lanham Act influenced the adult outcome distribution for treated relative to comparison cohorts. In particular, I rely on quantile regression methods developed by Firpo, Fortin, and Lemieux (2009) to assess whether the child-care program differentially influenced adult earnings at different points in the distribution.

The method proposed by Firpo et al. (2009) provides estimates of the effect of changes in a given right-hand-side variable on the unconditional quantiles of the left-hand-side variable. "Unconditional" in this context means the integration of the conditional outcome distribution, given a set of explanatory variables, over the distribution of those explanatory variables. The method consists of estimating a regression of the recentered influence function (RIF)—which characterizes the influence of each observation in the data on some distributional statistic—of the outcome variable on one or more explanatory variables. For a given quantile, the RIF, RIF(Y, q_τ), reflects $(\tau - \mathbf{I}\{Y \leq q_\tau\}) / f_Y(q_\tau)$, where $\mathbf{I}\{\cdot\}$ in an indicator function, $f_Y(\cdot)$ is the density of the unconditional distribution of the outcome variable, and q_τ is the population in the τth quantile of the outcome distribution. Because of $\mathbf{I}\{\cdot\}$, the RIF is a binary variable that equals $[-(1 - \tau)/f_Y(q_\tau)]$ when the observed value of the outcome is below the τth quantile, and it equals $[\tau / f_Y(q_\tau)]$ when the value is above the τth quantile. The conditional expectation of the outcome is specified as $E[\mathrm{RIF}(Y; q_\tau) \mid X]$, which is called the unconditional quantile regression. The conditional expectation at each quantile, τ, referred to as the DD quantile treatment effect (QTE), is straightforwardly generated using an OLS estimator.

As discussed in Havnes and Mogstad (2015), the DD QTE requires a slightly different identification assumption from that of the standard DD model. This stems from the fact that the counterfactual in the DD QTE is the change among comparison cohorts in the distribution of earnings across those born in low- and high-spending states. Therefore, identification of the various QTE parameters assumes that the change in population shares, q_τ, between treated and comparison cohorts at each quantile of earnings would be the same in low- and high-spending states in the absence of the Lanham Act. The reported DD QTE estimates can be interpreted as the difference in the distribution of earnings at a given quantile for treated individuals born in low- and high-spending states, relative to the difference for those in the comparison cohorts.

Figure 3, parts *A*, *B*, and *C*, present the DD QTE results for the 1970–90 Censuses, respectively. The outcome variable in these analyses is restricted to the log of annual earnings. Each dot depicts a QTE estimate on treated × lanham at the τth quantile, with the circles indicating that the estimate is statistically significant at the 10% level. The standard errors are bootstrapped with 500 replications, where random samples are drawn with replacement. Generally speaking, the figures reveal substantial heterogeneity in the impact of the Lanham Act. The largest positive effects are consistently found

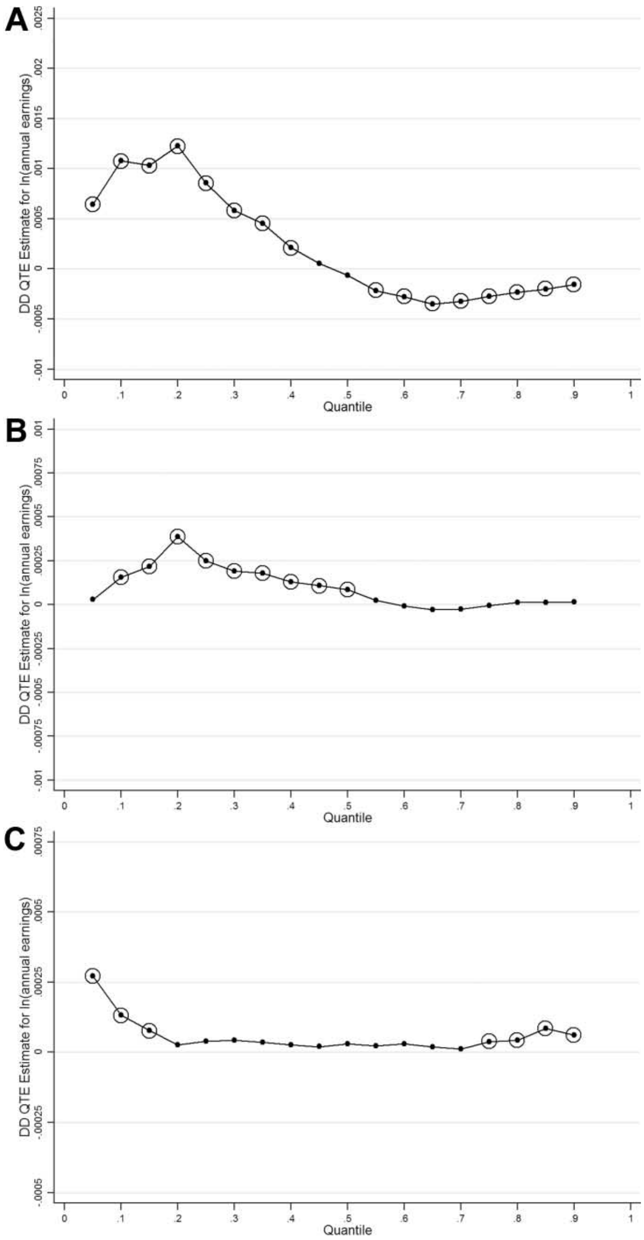

FIG. 3.—Difference-in-differences (DD) quantile treatment effect (QTE) estimates on annual earnings: *A*, 1970 US Decennial Census; *B*, 1980 US Decennial Census; *C*, 1990 US Decennial Census. Circles indicate that a given QTE estimate is statistically significant at the .10 level.

at the bottom of the earnings distribution, with the QTEs in each Census peaking somewhere in the lowest quartile of earnings. The QTEs then drop precipitously before flattening out throughout the remainder of the distribution. Indeed, the effects hover close to zero starting around the 50th percentile of earnings. Together, the pattern of results suggests that universal child care compressed the adult earnings distribution, relative to the counterfactual distribution, by increasing the earnings of the lowest earners substantially more than their higher-earning counterparts.

There are a few other noteworthy observations about the DD QTE estimates. First, it appears that most of the earnings compression occurred in the 1970 Census, when treated adults are in the early stages of their prime working years. This can be seen by the larger earnings QTEs in the bottom quartile of the 1970 distribution, relative to the QTEs in the same part of the 1980 and 1990 distributions. Greater compression is also evidenced by the slower flattening out of the QTEs in the 1970 Census. These earnings effects do not flatten out until approximately the 50th percentile; the QTEs in the 1990 Census are fairly steady beginning in the 20th percentile. Second, the QTE estimates for 1970 imply that the Lanham Act had negative effects on adult earnings starting at the 60th percentile. Such results are consistent with Havnes and Mogstad's (2015) QTE estimates for the Norwegian universal child care system, and they conform with recent US-based studies of pre-kindergarten programs (Cascio and Schanzenbach 2013) and nonparental child care arrangements (Herbst 2013), which uncover negative test score effects on children from more advantaged families.

F. Mechanisms

The final set of analyses examines the mechanisms driving the positive long-run impact of the Lanham Act. This discussion focuses on the counterfactual mode of child care, families' material resources, children's long-run educational attainment and health, and mothers' fertility response. As shown in table 7, I allow for the possibility that these mechanisms operate in the short run and the long run. Specifically, panel A studies families' short-run responses using the 1940 and 1950 Censuses, while panel B examines long-run responses using the 1940 and 1960 Censuses. In both cases, the estimates are derived from the DDD model outlined in equation (1). Panel C explores the long-run impact of the Lanham Act on children's educational attainment and health using the 1970 Census and the DD framework outlined in equation (2).

The first mechanism I consider is the counterfactual mode of child care. Two issues are central to this mechanism: whether children exposed to the Lanham Act would have been in maternal care or informal (i.e., relative or neighbor) arrangements in the absence of reform and whether the Lanham Act centers were of higher quality than the counterfactual mode of care. An-

Table 7
Short- and Long-Run Mechanisms

	Mean (1)	Observations (2)	DDD/DD Estimate (3)
A. Short-run mechanisms using the 1940 and 1950 Censuses:			
Part-time employment	.048	332,923	.00017**
	(.213)		(.00007)
Full-time employment	.120	423,153	.00099***
	(.325)		(.00029)
Ln(earnings last year)	8.801	92,817	.00073***
	(1.020)		(.00026)
Number of own children	2.363	247,809	−.00732***
	(1.491)		(.00260)
B. Long-run mechanisms using the 1940 and 1960 Censuses:			
Employed	.176	292,302	.00076***
	(.381)		(.00023)
Part-time employment	.047	206,196	.00017***
	(.211)		(.00006)
Full-time employment	.122	268,392	.00073***
	(.328)		(.00023)
Ln(earnings last year)	8.881	116,350	.00049**
	(.996)		(.00019)
C. Children's educational attainment and health using the 1970 Census:			
High school drop-out	.214	456,070	−.00021***
	(.410)		(.00005)
Bachelor's degree	.070	456,070	.00019***
	(.255)		(.00006)
Work-limiting disability	.062	456,070	−.00003
	(.241)		(.00002)

SOURCE.—Author's analysis of the 1940 and 1950 US Decennial Censuses (panel A), 1940 and 1960 US Decennial Censuses (panel B), and 1970 US Decennial Census (panel C) (IPUMS).

NOTE.—For panels A and B, col. 1 shows the outcome mean for the treated group in the pre-reform period; for panel C, col. 1 shows the outcome mean for the comparison cohorts. Column 2 shows the number of observations in the DDD or DD regressions, whose estimates are displayed in col. 3. Specifically, panels A and B show the DDD estimates from eq. (1), while panel C shows the DD estimates from eq. 2. Standard errors (in parentheses) are adjusted for state-level clustering.

** $p < .05$.
*** $p < .01$.

ecdotal evidence indicates that the Lanham Act induced a shift away from maternal care rather than informal arrangements. Prior to the war, mothers were expected to serve as the primary caretakers of preschool-aged children. Such cultural expectations were reinforced by the strong political opposition to virtually any form of nonparental care—formal or informal—that deviated from this norm (Berry 1993). Even after the war began, women with young children were allowed to work only after it became clear that the pool of childless women was insufficient for increasing war production.

The empirical evidence also supports the notion that mothers provided the counterfactual mode of care: the war in and of itself increased maternal employment in the short run (Goldin 1991) and the long run (Acemoglu et al. 2004; Goldin and Olivetti 2013), and the DDD evidence discussed earlier reveals that the Lanham Act increased maternal employment independently of the war's effect. Thus, the positive long-run impact of the child care program likely operates via a reduction in maternal care rather than informal care.

Regarding the relative quality of Lanham Act centers, the evidence suggests that children were exposed to child care of fairly high (if variable) quality. As previously stated, nurseries generally followed the recommended child-teacher ratio of 10-to-1, and most children engaged in various educational activities, were given time to complete homework and engage in extracurricular activities, and were provided with meals. Koshuk (1947) indicates that the centers in California—which served the largest number of children—provided very high quality services. Particularly noteworthy is that these centers either hired certified schoolteachers or provided on-site training and college credit to assist nonteachers. In addition, parents were pleased with the Lanham Act centers. Overwhelming majorities felt that their child progressed intellectually and socially and stated that their experience created a positive impression of early childhood education (Koshuk 1947).

Panels A and B empirically examine other mechanisms related to maternal employment and family material resources. Recall that the DDD analysis discussed earlier shows that the Lanham Act increased maternal employment in 1950—7 years after the program was terminated. Panel A reveals that the increase in employment was driven primarily by shifts into full-time work and to a lesser extent into part-time work. Given that these mothers became firmly rooted in the labor force, it is not surprising that the Lanham Act increased their earnings as well.[24] Panel B takes the analysis a step further by examining mothers' long-run labor market outcomes using the 1960 Census—17 years after the Lanham Act was discontinued. The same patterns emerge: treated mothers are more likely to be employed, are more likely to be working full-time, and have higher earnings. These results are consistent with the finding that the war in and of itself generated a long-run labor market response (Acemoglu et al. 2004; Goldin and Olivetti 2013). Thus, one set of mechanisms driving children's positive outcomes is the medium- and long-run labor market attachment of mothers and the resulting improvement in families' material resources.

[24] Although it would be beneficial to have data on total family income (as opposed to mothers' earnings), this information is not included in the 1940 Census. The only income variable found in the 1940, 1950, and 1960 Censuses is own wage and salary income.

Another set of mechanisms focuses on improvements in educational attainment and health. To examine these channels, I use the 1970 Census, which captures treated cohorts when they were ages 24–39. I examine two outcomes related to educational attainment: binary indicators for whether a given individual was a high school drop-out and whether the individual obtained at least a college degree. As shown in panel C, the DD estimates reveal that the Lanham Act produced broad improvements in educational attainment by lowering the high school drop-out rate and increasing the college completion rate. Similar results are found in the 1980 and 1990 Censuses' samples, indicating that these benefits persisted throughout adulthood. To examine the health effects of the Lanham Act, I use a Census item capturing the presence of physical or mental conditions that limit or prevent work. As shown in panel C, increases in Lanham Act spending reduced the incidence of work-related disabilities. Again, these health benefits persisted throughout adulthood.

The final mechanism focuses on the fertility response to the Lanham Act. Given that the child care program increased maternal labor supply in the medium run and the long run, it is possible that treated mothers reduced their fertility. A reduction in fertility could improve child outcomes by increasing parental time and resource investments, as predicted by some models of fertility (e.g., Becker and Lewis 1973). This implies that the DD estimates are larger than would be the case without such a fertility response. Conversely, if the Lanham Act reduced fertility rates, the opportunities for positive peer effects are diminished. This suggests that the DD estimates are smaller than would be the case if fertility rates were higher. To investigate this channel, I estimate a DDD model whose outcome is the number of own children in the household, using the 1940 and 1950 Censuses' sample.[25] Panel A shows that the Lanham Act had a small effect on fertility: the estimate implies that a move from the bottom to the top quartile of the Lanham

[25] Given that the fertility analysis uses the 1940 and 1950 Censuses' sample, the DDD coefficient provides evidence on the short- to medium-run fertility response. Although it would be useful to extend the analysis to the 1940 and 1960 Censuses' sample—to look at the long-run response—there is a significant complication that precludes such an analysis. In this sample, treated women are defined as those whose youngest child residing in the household is aged 14–29, while comparison women are those whose youngest child is age 30–34. Given that the outcome is the number of own children in the household, this variable is likely to be a poor approximation of the total number of children ever born to mothers in the treated and comparison groups. This is because the age restriction is so high for both groups that large numbers of age-eligible children no longer reside in the home. Those who remain may not be representative of the total population of 14–29 and 30–34 year-olds. In addition, it is possible that a large number of children born after the Lanham Act was terminated (but in response to the program) are not counted in the outcome variable because they too no longer reside in the home.

Act spending distribution would lead to 0.32 fewer children. This estimate is 14% of the pre-reform mean for the treated group. Therefore, it is unlikely that endogenous fertility threatens the causal interpretation of the results.

VI. Conclusion

The goal of this paper is to provide the first comprehensive analysis of the Lanham Act—a near-universal child care program operating throughout the United States during World War II. Although results from this analysis are likely to be of interest to a broad group of historians and economists as well as child care analysts, such results may also shed light on the contemporary policy push to enact universal child care programs. Indeed, missing from the debate, especially in the United States context, is credible evidence on the long-run impact of scaled-up, publicly funded, and universally accessible early care and education programs. The Lanham Act provides a potentially useful laboratory for assessing the promise of such programs.

Several noteworthy results stem from this analysis. First, I find that the Lanham Act increased maternal employment several years after the program was dismantled. Importantly, it generated approximately equal-sized increases in employment across a heterogeneous group of mothers. The second noteworthy result is that the Lanham Act had positive long-run effects on outcomes related to labor market behavior and program participation. Specifically, children exposed to the program were more likely to be employed, to have higher earnings, and to be less likely to receive cash assistance as adults. In addition, the benefits of the program accrued largely to the most economically disadvantaged individuals; in contrast, the program had neutral or even small negative effects on more advantaged adults. Finally, I show that the impacts operated through early-life increases in household income and long-run improvements in educational attainment and health.

On the one hand, it may be wise to interpret these results cautiously. As previously discussed, the Lanham Act was implemented during a unique period—amid a national emergency—in which traditional views about women's work and institutional child care were challenged and ultimately toppled. Such profound cultural shifts likely aided in sustaining the long-run employment effect of the program. This interpretation is consistent with the historical record on the Lanham Act (e.g., Fousekis 2011), and it accords with the characterization of the war as a "watershed" event for American women (Goldin and Olivetti 2013). One concern, then, is that the Lanham Act's antiquity and distinctive circumstances make it an inappropriate policy laboratory for considering the long-run impact of contemporary child care programs.

However, by definition, universal child care programs—including modern ones—are the product of unique political forces that, in turn, reshape the cultural and economic landscape.[26] Thus, the relevant question is, which characteristics of the Lanham Act account for its apparent success? In other words, are there lessons to be drawn from this experience that may inform the design of contemporary child care programs?

The first lesson is that while the program's primary aim was to increase women's employment, it did not come at the expense of lower child care quality or poorer outcomes for children. For example, the program's recommended child-to-teacher ratio (10 : 1) was more stringent than that in many states today. In addition, that the centers hired professional teachers and provided a university education to nonteachers suggests that the Lanham Act workforce was well prepared to handle children's developmental needs. Second, that the program had its largest impact on disadvantaged adults suggests that universal programs may be as effective as targeted ones. Indeed, one of the hypothesized mechanisms through which universal pre-K programs influence poor children is the social and intellectual engagement with their wealthy peers. However, as this and other studies show, such gains from universalism may come at a cost of worse outcomes for advantaged children. Third, the program generally shied away from serving infants and toddlers, for whom there were concerns about the consequences of prolonged separation from the mother. This may have been a fortuitous design feature in light of the empirical evidence that very young children in nonparental arrangements have worse outcomes in the short run than their peers in maternal care (Bernal and Keane 2011; Herbst 2013). Finally, the Lanham Act was a success because it received support from a broad coalition of parents, education and women's rights advocates, and employers. Each group was committed to its success because something larger was at stake: the nation's involvement in the war. Indeed, the rhetoric surrounding the Lanham Act—that expanding child care was a "patriotic" and "win-the-war" strategy—explicitly linked the need for child care with the nation's success on the battlefield.

[26] Quebec's universal child care program provides an illustration. Enacted in 1997, the program was one component in a large package of family-friendly policies that was unprecedented in its scope and cost. Government spending on child care grew from $288 million in 1997 to $2.2 billion in 2012, as the participation rate ballooned from 18% to 75%. In addition, the program fundamentally altered the supply of child care in Quebec, crowding out informal care and producing a threefold increase in the number of regulated child care spaces (Haeck, Lefebvre, and Merrigan 2015). The program's economic impact is large: in 2008 alone, an additional 70,000 mothers were employed, and Quebec's GDP grew by 1.7% because of universal child care (Fortin, Godbout, and St. Cerny 2012).

Appendix

Table A1
Difference-in-Differences (DD) Estimates from the Matched Case Control Analysis

DD Variable	Expenditures (1)	Comparison States (2)	Observations (3)	DD Estimate (4)
CA × treated	$264	AR, ID, NJ	43,945	.057***
				(.020)
WA × treated	$251	ME, ND, NH, NM NV, PA, SD, WI	60,498	.077***
				(.023)
OR × treated	$162	ID, NJ, NV	19,145	.055**
				(.022)
FL × treated	$141	AR, ID, NJ	29,530	.043***
				(.014)
AZ × treated	$140	AR, ID, KY, NJ, NV, PA, SD, WI, WV	89,109	.055**
				(.023)
GA × treated	$139	AR, ID, NJ, WV	42,631	.032**
				(.015)
UT × treated	$129	All but NH and NJ	84,211	.061**
				(.022)
DC × treated	$106	AR, ID, NJ	25,042	.021
				(.023)
DE × treated	$105	All but NJ	84,034	.089**
				(.036)
SC × treated	$100	ID, NJ	22,967	.054***
				(.016)
MS × treated	$81	AR, ID, NJ	30,914	.103***
				(.017)
AL × treated	$75	AR, ID, NJ, NV, WV	42,425	.026
				(.022)

Source.—Author's analysis of the 1970 US Decennial Census (IPUMS).

Note.—High-spending states are Alabama, Mississippi, South Carolina, Delaware, District of Columbia, Utah, Georgia, Arizona, Florida, Oregon, Washington, and California. Low-spending states are New Mexico, Idaho, West Virginia, North Dakota, Pennsylvania, Maine, Nevada, South Dakota, Arkansas, Kentucky, Wisconsin, New Hampshire, and New Jersey. Column 1 displays the Lanham Act expenditure amount (per child ages 0–12) for a given high-spending state. Column 2 displays the low-spending states included in the comparison group. The DD estimate and its standard error (in parentheses) are shown in col. 4. Standard errors are adjusted for birth-year clustering. The number of observations in the DD analysis is shown in col. 3. All models include the demographic controls outlined in eq. (2).

* $p < .10$.
** $p < .05$.
*** $p < .01$.

Table A2
Difference-in-Differences (DD) Estimates from the Placebo Reform

Variable	(1)	(2)	(3)
Treated × lanham	.00032**	.00019***	.00020***
	(.00013)	(.00005)	(.00006)
Placebo × lanham	9.18e-06	.00005	9.40e-06
	(.00010)	(.00005)	(.00004)

Table A2 (*Continued*)

Variable	(1)	(2)	(3)
F-statistic for Ho: $DD^{treated} = DD^{placebo}$	11.08	14.53	16.74
p-value	(.002)	(.000)	(.000)
Census sample	1970	1980	1990
Treated cohorts' age range	24–39	34–49	44–59
Placebo cohorts' age range	20–21	30–31	40–41
Comparison cohort's age	19	29	39

Source.—Author's analysis of the 1970, 1980, and 1990 US Decennial Censuses (IPUMS).
Note.—Displayed in each cell is the coefficient on treated × lanham and placebo × lanham and the standard errors (in parentheses), which are adjusted for state-level clustering. Results presented in col. 1 come from the 1970 Census, those in col. 2 are from the 1980 Census, and those in col. 3 are from the 1990 Census. All models include the full set of controls outlined in eq. (2).
* $p < .10$.
** $p < .05$.
*** $p < .01$.

Table A3
Additional Robustness Checks on the Main Difference-in-Diffrences (DD) Estimates

Specification	(1)	(2)	(3)
State-specific time trends	.00030***	.00011***	.00006*
	(.00007)	(.00004)	(.00004)
Alternative comparison group	.00028***	.00010*	.00025***
	(.00008)	(.00005)	(.00007)
Census sample	1970	1980	1990

Source.—Author's analysis of the 1970, 1980, and 1990 US Decennial Censuses (IPUMS).
Note.—Displayed in each cell is the coefficient on treated × lanham and its standard error (in parentheses), which is adjusted for state-level clustering. Results presented in col. 1 come from the 1970 Census, those in col. 2 are from the 1980 Census, and those in col. 3 are from the 1990 Census. The models in the first row include state-specific linear time trends in addition to all of the controls in the baseline model. The models in the second row use an alternative comparison group drawn from a set of pre-reform cohorts who were too old to have ever been age-eligible to receive child care under the Lanham Act.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

References

Acemoglu, Daron, David Autor, and David Lyle. 2004. Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of Political Economy* 112, no. 3:497–551.

Anderson, Karen. 1981. *Wartime women: Sex roles, family relations, and the status of women during World War II*. Westport, CT: Greenwood.

Anderson, Michael. 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association* 103, no. 484:1481–95.

Anderson, Patricia, and Phillip Levine. 2000. Child care and mothers' employment decisions. In *Finding jobs: Work and welfare reform*, ed. David Card and Rebecca Blank. New York: Russell Sage Foundation.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. Universal childcare, maternal labor supply, and family well-being. *Journal of Political Economy* 116, no. 4:709–45.

Barnett, W. Steven, Kwanghee Jung, Min-Jong Youn, and Ellen Frede. 2013. Abbott Preschool Program Longitudinal Effects Study: Fifth Grade Follow-up. NIEER Working Paper, National Institute for Early Education Research, Rutgers University.

Baum, Charles, and Christopher Ruhm. 2014. The effects of paid family leave in California on labor market outcomes. NBER Working Paper no. 19741, National Bureau of Economic Research, Cambridge, MA.

Becker, Gary, and Greg Lewis. 1973. On the interaction between the quantity and quality of children. *Journal of Political Economy* 81 (suppl.): S279–S288.

Bernal, Raquel, and Michael Keane. 2011. Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics* 29, no. 3:459–512.

Berry, Mary. 1993. *The politics of parenthood*. New York: Penguin.

Black, Sandra, Paul Devereux, Katrine Loken, and Kjell Salvanes. 2014. Care or cash? The effect of child care subsidies on student performance. *Review of Economics and Statistics* 96, no. 5:824–37.

Blau, David, and Erdal Tekin. 2007. The determinants and consequences of child care subsidies for single mothers in the USA. *Journal of Population Economics* 20, no. 4:719–41.

Cascio, Elizabeth. 2009a. Do investments in universal early education pay off? Long-term effects of introducing kindergartens into public schools. NBER Working Paper no. 14951, National Bureau of Economic Research, Cambridge, MA.

———. 2009b. Maternal labor supply and the introduction of kindergartens into American public schools. *Journal of Human Resources* 44, no. 1:140–70.

Cascio, Elizabeth, and Diane Schanzenbach. 2013. The impacts of expanding access to high-quality preschool education. Conference draft for the 2013 Brookings Panel on Economic Activity, Brookings Institution, Washington, DC.

Connelly, Rachel, and Jean Kimmel. 2003. The effect of child care costs on the employment and welfare recipiency of single mothers. *Southern Economic Journal* 69, no. 3:498–519.

Conti, Gabriella, James Heckman, and Rodrigo Pinto. 2015. The effects of two influential early childhood interventions on health and health behaviors. NBER Working Paper no. 21454, National Bureau of Economic Research, Cambridge, MA.

Deming, David. 2009. Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics* 1, no. 3:111–34.
Dratch, Howard. 1974. The politics of child care in the 1940s. *Science and Society* 38, no. 2:167–204.
Federal Works Agency. Various years. *Annual report of the Federal Works Agency*. Washington, DC: Superintendent of Documents, Government Printing Office.
Firpo, Sergio, Nicole Fortin, and Thomas Lemieux. 2009. Unconditional quantile regressions. *Econometrica* 77, no. 3:953–73.
Fitzpatrick, Maria. 2008. Starting school at four: The effect of universal prekindergarten on children's academic achievement. *B. E. Journal of Economic Analysis and Policy* 8, no. 1, article 46.
———. 2010. Preschoolers enrolled and mothers at work? The effects of universal prekindergarten. *Journal of Labor Economics* 28, no. 1:51–85.
———. 2012. Revising our thinking about the relationship between maternal labor supply and preschool. *Journal of Human Resources* 47, no. 3: 583–612.
Fortin, Pierre, Luc Godbout, and Suzie St. Cerny. 2012. Impact of Quebec's universal low fee childcare program on female labour force participation, domestic income, and government budgets. Working Paper no. 2012–02, University of Sherbrooke.
Fousekis, Natalie. 2011. *Demanding child care: Women's activism and the politics of welfare, 1940–1971*. Champaign: University of Illinois Press.
Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. Longer-term effects of Head Start. *American Economic Review* 92, no. 4:999–1012.
Gelbach, Jonah. 2002. Public schooling for young children and maternal labor supply. *American Economic Review* 92, no. 1:307–22.
Goldin, Claudia. 1991. The role of World War II in the rise of women's employment. *American Economic Review* 81, no. 4:741–56.
Goldin, Claudia, and Claudia Olivetti. 2013. Shocking labor supply: A reassessment of the role of World War II on women's labor supply. *American Economic Review* 103, no. 3:257–62.
Gormley, William, and Ted Gayer. 2005. Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. *Journal of Human Resources* 40, no. 3:533–58.
Gupta, Nabanita Datta, and Marianne Simonsen. 2010. Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics* 94, no. 1:30–43.
Haeck, Catherine, Pierre Lefebvre, and Philip Merrigan. 2015. Canadian evidence on ten years of universal preschool policies: The good and the bad. *Labour Economics* 36:137–57.

Havnes, Tarjei and Magne Mogstad. 2011a. Money for nothing? Universal child care and maternal employment. *Journal of Public Economics* 95, no. 11:1455–65.

———. 2011b. No Child Left Behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy* 3, no. 2:97–129.

———. 2015. Is universal child care leveling the playing field? *Journal of Public Economics* 127:100–114.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter Savalyev, and Adam Yavitz. 2010. Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics* 1, no. 5095:1–46.

Herbst, Chris M. 2010. The labor supply effects of child care costs and wages in the presence of subsidies and the Earned Income Tax Credit. *Review of Economics of the Household* 8, no. 2:199–230.

———. 2013. The impact of non-parental child care on child development: Evidence from the summer participation "dip." *Journal of Public Economics* 105 (September): 86–105.

Herbst, Chris M., and Erdal Tekin. 2012. The geographic accessibility of child care subsidies and evidence on the impact of subsidy receipt on childhood obesity. *Journal of Urban Economics* 71, no. 1:37–52.

———. 2016. The impact of child care subsidies on child well-being: Evidence from geographic variation in the distance to social service agencies. *Journal of Policy Analysis and Management* 35, no. 1:94–116.

Johnson, Marilynn. 1993. *The second gold rush: Oakland and the East Bay in World War II*. Berkeley: University of California Press.

Kerr, Virginia. 1973. One step forward—two steps back: Child care's long American history. In *Child care: Who cares?* ed. Pamela A. Roby. New York: Basic Books.

Koshuk, Ruth. 1947. Developmental records of 500 nursery school children. *Journal of Experimental Education* 16, no. 2:134–48.

Lefebvre, Pierre, and Philip Merrigan. 2008. Child care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics* 26, no. 3:519–48.

Ludwig, Jens, and Douglas Miller. 2007. Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics* 122, no. 1:159–208.

McCormick, M. C., Jeanne Brooks-Gunn, S. L. Buka, J. Goldman, J. Yu, and M. P. Salganik. 2006. Early intervention in low birth weight premature infants: Results at 18 years of age for the Infant Health and Development Program. *Pediatrics* 117, no. 3:771–80.

NIEER (National Institute for Early Education Research). 2012. The state of preschool, 2012. National Institute for Early Education Research, Rutgers University.

Riley, Susan. 1994. Caring for Rosie's children: Federal child care policies in the World War II era. *Polity* 26, no. 4:655–75.

Ruggles, Steven Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]. University of Minnesota, Minneapolis.

Stoltzfus, Emilie. 2000. Childcare: The federal role during World War II. US Congressional Research Service, RS20615.

Tank, Robert. 1980. Young children, families, and society in America since the 1820s: The evolution of health, education, and child care programs for preschool children. PhD diss., University of Michigan, Ann Arbor.

Tekin, Erdal. 2005. Child care subsidy receipt, employment, and child care choices of single mothers. *Economics Letters* 89, no. 1:1–6.

———. 2007a. Childcare subsidies, wages, and employment of single mothers. *Journal of Human Resources* 42, no. 2:453–87.

———. 2007b. Single mothers working at night: Standard work and child care subsidies. *Economic Inquiry* 45, no. 2:233–50.

Tuttle, William, Jr. 1995. Rosie the Riveter and her latchkey children: What Americans can learn about child day care from the Second World War. *Child Welfare* 74, no. 1:92–114.

US House of Representatives. 1945. To amend Section 204, Title II, of the Act to Expedite Provision of Housing. Committee on Public Buildings and Grounds. 79th Congress, 1st session, May 16 and 22.

US Senate. 1943a. Wartime care and protection of children of employed mothers. Committee on Education and Labor. 78th Congress, 1st session, June 8.

———. 1943b. Wartime health and education. Committee on Education and Labor, Subcommittee on Wartime Health and Education. 78th Congress, December 2.

Weiland, Christina, and Hirokazu Yoshikawa. 2013. Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development* 84, no. 6:2112–30.

Youcha, Geraldine. 1995. *Minding the children: Child care in America from colonial times to the present*. New York: Scribner.