The Effects of Infant Daycare on Later-in-Life Employment Outcomes

David M. Zimmer*
Western Kentucky University

February 15, 2021

Abstract

This paper first develops a theoretical model of the returns to daycare, which attempts to show that families enroll their children into daycare based on its expected benefits and costs. The paper then moves to an empirical study that seeks to determine whether enrolling an infant in daycare affects his or her later-in-life employment and income. To identify the causal effect of interest, the econometric approach employs a recently-developed panel estimator that accommodates the dynamic nature of later-in-life employment and income, while also controlling for unobserved heterogeneity. The main finding is that, although infant daycare enrollment appears to correlate with positive later-in-life outcomes, that link largely stems from parents of children who would have experienced positive outcomes anyway enrolling their children in daycare.

JEL Codes: C23; D13

Keywords: fixed effects; dynamic; childcare

*david.zimmer@wku.edu; Grise Hall 426, Western Kentucky University, Bowling Green, KY 42101

1 Introduction

During the summer of 2019, the New York Times reported on a study by the National Institute of Child Health and Human Development (NICHD) purporting to show long-term negative effects of infant daycare ("How to End the Childcare Crisis." New York Times, 5/25/2019.) Other widely-followed media outlets and think tanks – including CBS, Livestrong.com, and the Heritage Foundation – further disseminated those findings. That media attention created much discussion and debate about our nation's childcare priorities.

The most oft-quoted reason for suspecting negative effects from daycare revolves around whether a child's time spent with daycare staff – which likely must look after multiple children – represents an inferior substitute for time spent with a parent (Long and Garduque, 1987). Another concern is that daycare might place children at increased risk for serious illness (Haskins and Kotch, 1986). On the other hand, another branch of research posits that being around other children under the supervision of trained professionals potentially fosters emotional and pro-social development (Gomajee et al., 2018). Moreover, if daycare facilitates parental labor force participation, as suggested by a relatively large strand of research (Heckman, 1974; Kimmel, 1998; Blau, 2000; Spiess and Wrohlick, 2008; Bick, 2010; Havnes and Mogstat, 2011; Domeij and Klein, 2013), then perhaps increased income leads to improved child outcomes.

But any study on the long-term effects of daycare must confront a crucial complication: We never observe the same infant *both enrolled and not enrolled* in daycare. The impossibility of observing that dichotomy could be sidestepped *if infants were randomly assigned to daycare.* But, of course, they are not. Infants usually enroll in daycare because their parents have some form of attachment to the labor force, with daycare serving the role as a daily babysitter. Moreover, enrolling infants in daycare requires households to have access to financial resources necessary to pay the (often steep) tuition fees. Thus, if parental employment and income correlate with child outcomes, as suggested by an enormous body of research, then perhaps those traits, rather than daycare itself, are the true channels affecting a child's later-in-life employment outcomes. Consequently, any study that simply compares daycare enrollees to non-enrollees quickly runs aground.

Perhaps for those reasons, existing research produces a wide range of conflicting findings. As an abbreviated sample of existing work, studies have found that daycare increases behavior problems (Belsky, 1999; Vandell, Burchinal, Friedman, and Brownell, 2001), while others have found that it decreases them (Denham and Burton, 1996; Field, 1991; Prodromidis, Lamb, Sternberg, Hwang, and Broberg, 1995). Some studies have found negative effects of daycare on cognitive development (Russell, 1999), while others have found positive effects (Spieker, Nelson, Petras, Jolley, and Barnard, 2003). Mostly owing to data limitations, existing studies almost exclusively explore the effects of daycare on relatively short-run outcomes, and, as a result, little information exists regarding longer-run impacts, such as those explored in this paper.

This paper first develops a theoretical model of the returns to daycare. The goal is to show that families enroll their children into daycare based on its expected benefits and costs. The paper then moves to an empirical study that seeks to determine whether enrolling an infant in daycare affects his or her later-in-life employment and income. But, as argued by Shpancer (2006), establishing a causal link from day care to socioeconomic outcomes is potentially muddled by the likely presence of confounding factors, with families that enroll their infants in daycare likely possessing unmeasured traits that also correlate with later-in-life outcomes. To identify the causal effects of interest, the econometric approach employs a recently-developed panel estimator, developed by Kripfganz and Schwarz (2019), that (1) accommodates the dynamic nature of later-in-life employment and income, and controls for unobserved heterogeneity.

The paper produces several findings. First, in terms of simple correlations, infant daycare enrollment associates with a 6 percent increase in later-in-life full time work, a 5 percent increase in the likelihood of being over 200 percent of the poverty line, and an 8.5 percent increase in later-in-life annual income. However, those associations vanish once one accounts for dynamic effects and person-specific unobserved heterogeneity. Thus, the main conclusion is that, although infant daycare enrollment appears to *correlate* with positive later-in-life outcomes, that link largely stems from parents of children who would have experienced positive outcomes anyway enrolling their children in daycare. Therefore, infant daycare appears *not* to exert any direct causal effects on later-in-life employment and income.

2 The Returns to Daycare

Borrowing from theoretical models of the effects of human capital formation developed by Becker (1967) and Card (1999), this section sketches a model that attempts to explain the causal effect of infant daycare on later-in-life outcomes. The model seeks to emphasize that the returns to daycare show differences across individuals, with those differences driven by heterogeneity in feelings toward, and expected benefits of, daycare. That heterogeneity, in turn, muddles the *observed* link between daycare enrollment and later-in-life outcomes.

Let $y(d)$ denote the average later-in-life outcome of interest resulting from infant daycare enrollment d. For ease of exposition, let y and d both represent *amounts*, so that each may be treated as continuous, even though their observed counterparts in the estimation sample might be discrete. The parents of child i seek to maximize the utility function

$$U(y(d), d) = \log y(d) - c(d)$$

where costs of daycare, represented by $c(d)$, are likely convex, reflected by $c'(d) > 0$ and $c''(d) > 0$. Convexity of costs would be likely if the marginal cost of more daycare rises by more than forgone wages for a parent staying home with the infant, perhaps due to credit market restrictions or taste factors (Becker, 1967). Card (1999) highlights that this utility setup generalizes to a discounted present value objective function, with subjects discounting future outcomes at a constant rate.

Differentiating the utility function and setting equal to zero, the optimal amount of daycare must satisfy

$$\frac{y'(d)}{y(d)} - c'(d) = 0. \tag{1}$$

As a simple way to allow heterogeneity in the returns to daycare, write the marginal return to daycare as

$$\frac{y'(d)}{y(d)} = a_i - k_1 d \tag{2}$$

where a_i represents differences in the economic benefits of daycare, and the negative sign attached to the non-negative constant k_1 captures (likely) concavity of the marginal return. Similarly, to allow heterogeneity in costs, write the marginal cost as

$$c'(d) = b_i + k_2 d \tag{3}$$

where b_i captures differences in cost – including psychological costs, like possible distaste for daycare – and the non-negative constant k_2 reflects (likely) convexity of costs.

Plugging (2) and (3) into (1) and solving for d gives optimal daycare,

$$d^* = \frac{a_i - b_i}{k_1 + k_2}. \tag{4}$$

This simple model equilibrium entails a distribution of costs and returns, because the terms a_i and b_i vary across children. This equation also highlights possible endogeneity of daycare enrollment with respect to future outcomes, thus muddling the observed link between the two. The source of that endogeneity is that, by equation (4), the terms a_i and b_i influence a family's level of daycare enrollment. But equations (2) and (3) show that those same terms affect (perceived) benefits and costs of daycare.

Diving further into endogeneity, let $\log y(d)$ be increasing and concave, implying that daycare contributes to utility positively, but at a diminishing rate (although, to be sure, ascertaining the shape of that marginal return is the main empirical focus of this paper). In that case, a family that expects its child to reap large benefits from daycare (i.e., a larger value of a_i) would see larger returns to daycare, because the marginal return $\frac{y'(d)}{y(d)}$ increases in a_i. And if the family's expectations end up being correct, then infants enrolled in daycare would enjoy higher later-in-life y. But that does *not* imply that children randomly assigned to daycare would reap similar benefits. Rather, it only means that children with similar values of a_i would reap comparable benefits.

On the cost side, a family that resides far from a daycare facility or that has a particular distaste for daycare (i.e., a larger value of b_i) would have higher marginal costs. Equation (4) shows that those higher costs reduce daycare enrollment. And if $y(d)$ is increasing, then less daycare leads to lower later-in-life y. But, again, that relation cannot be interpreted as causal. Rather, the relationship indicates that children who face higher marginal costs of daycare experience worse later-in-life outcomes.

In sum, not only does this theoretical framework produce ambiguous conclusions about the causal effects of daycare. It also suggests that child- or family-specific heterogeneity in the (expected) returns to and costs of daycare precludes a simple comparison of daycare enrollees to non-enrollees. Thus the topic becomes an empirical investigation aimed at identifying causality. Section (4) outlines a relatively new estimator that attempts to tease out that causal information from observation data.

3 Data

The main obstacle in linking daycare enrollment to later-in-life outcomes comes from the relative paucity of micro surveys that contain information on a person's childcare arrangements during his or her early life. Further, because later-in-life outcomes likely show dynamic patterns, those same micro surveys need to track those later-in-life outcomes over time. Perhaps the only nationally-representative survey that meets those conditions is the 1997 National Longitudinal Survey of Youth (NLSY97).

Sponsored by the U.S. Bureau of Labor Statistics, the NLSY97 provides a nationally-representative survey of approximately 9,000 youths born during the years 1980 and 1984 and living in the United States. Those youths were 12 to 16 years old as of the survey's inception in December 31, 1996. Respondents were interviewed annually until 2011, and biennially after that.

The key treatment variable in this study is a binary indicator for whether, during the respondent's first year of life, he or she spent at least 20 hours per week cared for by someone other than his or her parents. That variable comes from the initial wave of the survey (1997) and, because it refers to the respondent's first year of life, does not vary thereafter. To reduce the possibility that such care happened through informal arrangements, the sample does not include subjects who listed non-parents (such as grandparents) as their primary infant care givers. For brevity, the remainder of this study refers to that variable as *daycare*. According to that definition, approximately 29 percent of the estimation sample enrolled in infant daycare, a number that aligns with national surveys of daycare enrollment (Glynn, 2012).

Three other variables – *female*, *black*, and *Hispanic* – also come from the initial wave of the survey (1997) and do not vary thereafter. A fourth variable – *age in 2008* – comes from the 2008 wave, and also is treated as time invariant. Table 1 reports sample means for those time-invariant measures, partitioned by infant daycare status. Blacks appear slightly more likely to have enrolled in daycare, while Hispanics appear less likely to have enrolled in daycare. Otherwise, neither age nor gender appears to differ significantly according to daycare status.

To correspond with "adulthood," all time-varying information comes from the 2008, 2009, 2010, 2011, 2013, 2015, and 2017 waves of the survey, during which respondents would have been between the ages 24-38. (Note that the periodicity of the survey switches from annual to biennial after 2011. Dropping the biennial years did not alter the main conclusions of this paper.) Table 2 shows means for four time-varying covariates that serve as controls in the methods described below. Of the four, only the presence of children differs significantly by daycare status, with daycare enrollees appearing less likely to have children later in life, although the magnitude of that difference appears to be quantitatively small.

Table 3 shows the three outcome variables of interest, which, like those listed in Table 2, vary over time. All three of those variables suggest statistically-significant, and relatively-sizable, positive associations with having been enrolled in infant daycare. However, as cautioned in the previous section, those associations cannot be interpreted as causal due to the likelihood that subjects who were enrolled in infant daycare possess unmeasured attributes that likely correlate with later-in-life outcomes.

The methods discussed in the following section aim to remove bias introduced by those unmeasured attributes. But those methods depend, in part, on the three outcome measures showing sufficient intra-person variation over time. Although no guidance exists on what constitutes “sufficient” variation, Table 4 reports coefficients of variation (within-person standard deviations divided by overall means) for those three measures, and also the other time-varying measures for comparison. Notably, the three outcome variables appear to show intra-person variation similar to other time-varying measures, particularly marital status, the presence of children, and the regional unemployment rate. The presence of such variation should allow the estimation approach to achieve reasonable precision for estimate key parameters of interest. (Note that regional housing starts show substantial variation, which should assist the identification strategy discussed in the following section.)

4 Methods

Let $y_{i,t}$ denote person i's later-in-life outcome in year t. Let that outcome be expressed linearly as

$$y_{i,t} = \lambda y_{i,t-1} + \mathbf{X}'_{it}\boldsymbol{\beta} + \gamma d_i + \alpha_i + \varepsilon_{it}. \tag{5}$$

The vector $\mathbf{X}_{it}$ includes the aforementioned time-varying controls: marital status, presence of children, regional unemployment rate, and regional percent change in housing starts. Those time-varying controls have estimable coefficients $\boldsymbol{\beta}$, and the term ε_{it} represents an error. The main explanatory variable of interest, d_i, indicates whether the person was enrolled in daycare during the first year of his or her life. The main parameter of interest, then, is the coefficient γ.

Two important elements in equation (1) deserve mention. First, the person's outcome state is *dynamic*, as $y_{i,t-1}$ appears on the right-hand side. (Employment and income tend to show strong serial persistence, thus necessitating a dynamic setup.) Second, the term α_i, treated here as a "fixed effect" in the econometrics sense of the term, captures all time-invariant attributes pertaining to person i that we *do not* observe. For example, why was person i enrolled (or not) in daycare as an infant? Were his parents employed? Were they financially secure? Or did they simply value whatever they thought daycare had to offer? Although we do not observe that information, due to their time invariance (once person i reaches adulthood), their influence is captured by α_i. (Importantly, those time-invariant unobserved factors may correlate with observed variables, including infant daycare enrollment.)

However, as presented, equation (1) presents two complications. First, including both $y_{i,t-1}$ and α_i on the right-hand side biases all coefficient estimates, a plague known as "Nickel Bias." The most famous solution to this problem, developed by Arellano and Bond (1991), is simple to implement, but somewhat fragile in practice. Therefore, many researchers prefer the more robust, yet computationally-taxing, method introduced by Hsiao et al. (2002).

Second, the main right-hand side variable of interest, d_i, does not vary over time, which means it perfectly correlates with the fixed effect α_i. Consequently, neither is separately identifiable without some modification. The classic solution to this problem, developed by Hausman and Taylor (1981), requires at least one time-varying element of $\mathbf{X}_{it}$ to remain uncorrelated with α_i.

Building upon earlier work by Hoeffler (2002), Cinyabuguma and Putterman (2011), and Pesaran and Zhou (2018), Kripfganz and Schwarz (2019) introduce a two-stage estimator that, in effect, blends the Hsiao et al. (2002) and Hausman and Taylor (1981) approaches. In the first stage, the time-invariant measure of daycare enrollment, d_i, is subsumed into the fixed effect, $\eta_i = \gamma d_i + \alpha_i$, which yields the first-stage model

$$y_{i,t} = \lambda y_{i,t-1} + \mathbf{X}_{it}'\boldsymbol{\beta} + \overline{\eta} + u_{it}$$

where

$$u_{it} = \eta_i - \overline{\eta} + \varepsilon_{it}.$$

Any root-N consistent estimator suffices in that first stage, in particular the Hsiao et al. (2002) approach. Then, the second stage uncovers the coefficient of the time-invariant daycare variable via the relationship

$$y_{i,t} - \widehat{\lambda} y_{i,t-1} - \mathbf{X}_{it}'\widehat{\boldsymbol{\beta}} = \gamma d_i + \upsilon_{i,t}$$

where

$$\upsilon_{i,t} = \alpha_i + \varepsilon_{it} - (\widehat{\lambda} - \lambda) y_{i,t-1} - \mathbf{X}'_{it}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}).$$

Kripfganz (2016) discusses a Stata algorithm that consistently estimates both stages, including adjusting the second-stage standard errors to account for first-stage statistical noise.

Similar to Hausman and Taylor (1981), identification of the two-stage estimator requires as least one time-varying element of $\mathbf{X}_{it}$ to remain uncorrelated with the fixed effect α_i. This paper uses two such measures: the current unemployment rate in the census region in which the person resides, and, similarly, the current percent change in housing starts in the person's region. Both plausibly relate to a person's current socioeconomic status. But more importantly, since they capture region-wide economic conditions, they plausibly remain uncorrelated with the fixed effect, which reflects *person-specific* heterogeneity.

The main threat to the validity of those two measures is the possibility that highly-motivated individuals – i.e., those likely to enjoy higher socioeconomics outcomes regardless of their infant daycare enrollment – might migrate to census regions with attractive economic conditions. Such migration patterns would impart correlation between those two measures and the fixed effect, since the region-specific measures would now, in part, reflect person-specific actions aimed at finding favorable economic conditions. Section 6 explores that potential threat to identification.

5 Results

Appendix Tables 1-3 present estimates of models that *do not* exploit the panel nature of the data, other than clustering the standard errors at the person level. Appendix 4 presents results from the two-step Kripfganz and Schwarz estimator that allows for both dynamics and fixed effects. Across all specifications, coefficients of the explanatory variables corroborate *a priori* expectations, so those are not elaborated upon here.

It is worth noting, however, that, as shown in the Appendix Tables, the measures of regional economic conditions, which serve as a source of identification, appear to affect subjects' outcomes in the expected directions. That is, positive growth in housing starts appears to correlate with a person's likelihood of working full-time, while higher unemployment rates associate with lower income and smaller likelihoods of being above 200 percent of the poverty line.

To distill the paper's main findings into one place, Table 5 presents coefficients of infant daycare enrollment. Estimates in the first row, derived from models that include only a constant, suggest that infant daycare enrollment associates with a 3.9 percentage point increase in later-in-life full time work. That increase corresponds to an approximate 6 percent increase, relative to the sample mean of full time work. Similarly, daycare correlates with a 3.7 percentage point increase in the likelihood of being over 200 percent of the poverty line, an approximate 5 percent boost relative to the sample mean. Finally, daycare corresponds to a marginally-significant (p value $= 0.11$) 8.5 percent increase in later-in-life income.

Those effects in the "Constant only" row, which are highly-likely to suffer contamination from omitted confounding factors, mostly correspond to the raw differences in sample means reported in Table 3. Curiously, the addition of controls, reported in the second row of Table 5, does not substantially alter those effects. In fact, the effect of daycare on later-in-life income *increases* to 10.6 percent, and becomes statistically significant.

The third row adds the lagged outcome. As reported in the Appendix Tables, the coefficients of those lagged outcomes are large in magnitude and highly-statistically significant, which points to serial persistence in those outcomes, and also attests to the importance of modeling those outcomes dynamically. Table 5 reveals that, after including those lags, the effects of daycare remain statistically significant, but shrink substantially in magnitude. Now, daycare boosts the likelihood of full-time employment by 1.5

percentage points, the probability of being above 200 percent of the poverty line by 1.9 percentage points, and later-in-life income by 5.2 percent.

The final row of Table 5 formally exploits the panel structure of the data by adding fixed effects, with estimates obtained using the Kripfganz and Schwarz two-step estimator. For all three outcomes, the effect of daycare becomes statistically-indistinguishable from zero. The interpretation is that, although infant daycare enrollment appears to *correlate* with positive later-in-life outcomes, that link largely stems from parents of children who would have experienced positive outcomes anyway enrolling their children in daycare.

6 A Check on Identification

Similar to the classic Hausman and Taylor (1981) approach it builds upon, the Kripfganz and Schwarz estimator rests on a key identifying assumption. That is, at least one time-varying control must remain uncorrelated with the person-specific fixed effect α_i. This paper uses two such time-varying controls: (1) the current unemployment rate in the census region in which the person resides and (2) the current percent change in housing starts in the person's region. Those two variables, while appearing to correlate with later-in-life outcomes in the expected directions, refer to *region-wide* economic conditions, and therefore plausibly do not relate to unobserved *individual-specific* traits.

The main threat to the validity of those two variables is the possibility that people choose to locate in parts of the country with favorable economic conditions. For example, a person might move to a locale with high-paying jobs, possibly elevating his or her income. But in that scenario, the regional economic situation correlates with the person's propensity to locate there, thus violating the identifying assumption.

To check whether such migrating behavior exists in sufficient quantity to pose a threat to the identification strategy, all models were re-estimated after adding a time-varying indicator for whether the person migrated to a different state since the previous survey year. Results, not presented in table form here, were nearly identical to those appearing in Table 5. Along similar lines, all models were re-estimated using a subsample of subjects who reported never having migrated between states at any time during the years 2008-2017. Again, results were nearly identical to those reported in Table 5.

Based on those (admittedly-informal) robustness checks, the regional measures of economic conditions seem to remain uncorrelated with unobserved person-specific contributors to later-in-life employment and income. Stately differently, migration owing to regional-wide economic conditions does not appear to exist in a sufficient amount to pose a threat to identification. Consequently, the regional economic measures appear to provide plausible sources of exogenous variation for the two-step estimator.

7 Conclusion

A study by the National Institute of Child Health and Human Development, and the subsequent media attention that ensued, sparked a debate during the summer of 2019 on the effects of daycare. But establishing the causal effects of daycare is difficult, due to lack of random assignment. Perhaps for that reason, existing studies produce a range of (often conflicting) findings.

Most existing studies explore the effects of daycare on relatively short-run outcomes, such as behavior, social development, progress through school, and standardized test scores. This paper, by contrast, investigates daycare's potential longer-run effects on employment and earnings. Linking early-life daycare enrollment to later-in-life outcomes has fairly steep data requirements, which fortunately are uniquely met by the 1997 National Longitudinal Survey of Youth.

This paper first presents a theoretical model of the returns to daycare, which aims to shed light on its ambiguous effects. The paper then presents an empirical panel data model that employs a recently-developed estimator that accounts for both dynamic effects and person-specific unobserved heterogeneity. Results show that, although infant daycare enrollment appears to correlate with improved later-in-life outcomes, those links largely stem from unobserved confounding factors. Once those are taken into account, infant daycare appears to exert no significant influence on later-in-life outcomes.

That finding does not imply that daycare offers zero societal benefits. As some of aforementioned studies reveal, daycare, despite its costs, potentially increases parental labor force attachment. Consequently, a full cost-benefit analysis of daycare cannot rely solely on its effects on children's eventual employment and income. Rather, such an analysis must incorporate larger general equilibrium considerations.

References

Arellano, M. and S. Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies*, 58, 277-297.

Becker, G. (1967). *Human Capital and the Personal Distribution of Income: An Analytical Approach.* Institute of Public Administration: Ann Arbor, MI.

Belsky, J. (1999). "Quantity of nonmaternal care and boys' problem behavior/adjustment at ages 3 and 5: exploring the mediating role of parenting." *Psychiatry*, 62, 1-20.

Bick, A. (2016). "The quantitative role of child care for female labor force participation and fertility." *Journal of the European Economic Association*, 14, 639-668.

Blau, D. (2000). "The production of quality in child-care centers: another look." *Applied Developmental Science*, 4, 136–149.

Card, D. (1999). "The causal effect of education on earnings." *Handbook of Labor Economics*, 3, 1801-1863.

Cinyabuguma, M. and L. Putterman (2011). "Sub-saharan growth surprises: being heterogeneous, inland and close to the equator does not slow growth within Africa." *Journal of African Economies*, 20, 217-262.

Denham, S and R. Burton (1996). "A social-emotional intervention for at-risk 4-year-olds." *Journal of School Psychology*, 34, 225-245.

Domeij, D. and P. Klein (2013). "Should day care be subsidized?" *Review of Economic Studies*, 80, 568-595.

Field, T. (1991). "Quality infant day-care and grade school behavior and performance." *Child Development*, 62, 863-870.

Glynn, S. (2012). "Child Care: Families Need More Help to Care for Their Children." Center for American Progress: Fact Sheet.

Gomajee, R., El-Khoury, F., Côté, S., Van Der Waerden, J., Pryor, L., and M. Melchior (2018). "Early childcare type predicts children's emotional and behavioural trajectories into middle childhood. Data from the EDEN mother–child cohort study." *Journal of Epidemiology and Community Health*, 72, 1033-1043.

Haskins, R. and J. Kotch (1986). "Day care and illness: evidence, costs, and public policy." *Pediatrics*, 77, 951-982.

Hausman, J. and W. Taylor (1981). "Panel data and unobservable individual effects." *Econometrica*, 49, 1377-1398.

Havnes, T. and M. Mogstad (2011). "No child left behind: subsidized child care and children's long-run outcomes." *American Economic Journal: Economic Policy*, 3, 97-129.

Heckman, J. (1974). "Effects of child-care programs on women's work effort." Journal of Political Economy, 82, S136-S163.

Hoeffler, A. (2002). "The augmented Solow model and the African growth debate." *Oxford Bulletin of Economics and Statistics*, 64, 135-158.

Hsiao, C., Pesaran, M., and A. Tahmiscioglu (2002). "Maximum likelihood estimation of fixed effects dynamic panel models covering short time periods." *Journal of Econometrics*, 109, 107-150.

Kimmel, J. (1998). "Child care costs as a barrier to employment for single and married mothers." *Review of Economics and Statistics*, 80, 287-299.

Kripfganz, S. (2016). "Quasi-maximum likelihood estimation of linear dynamic short-T panel data models." *Stata Journal*, 16, 1013-1038.

Kripfganz, S. and C. Schwarz (2019). "Estimation of linear dynamic panel data models with time-invariant regressors." *Journal of Applied Econometrics*, 34, 526-546.

Long, F. and L. Garduque (1987). "Continuity between home and family day care: Caregivers' and mothers' perceptions and children's social experiences." *Advances in Applied Developmental Psychology*, 2, 69–90.

Pesaran, M. and Q. Zhou (2018). "Estimation of time-invariant effects in static panel data models." *Econometric Reviews*, 37, 1137-1171.

Prodromidis, M., Lamb, M., Sternberg, K., Hwang, C., and A. Broberg (1995). "Aggression and noncompliance among Swedish children in centre-based care, family day care, and home care." *International Journal of Behavioral Development*, 18, 43-62.

Russell, C. (1999). "A meta-analysis of published research on the effects of nonmaternal care on child development." Dissertation Abstracts International Section A: Humanities & Social Sciences, 59(9-A), 3362.

Shpancer, N. (2006). "The effects of daycare: persistent questions, elusive answers." *Early Childhood Research Quarterly*, 21, 227-237.

Spieker, S., Nelson, D., Petras, A., Jolley, S., and K. Barnard (2003). "Joint influence of child care and infant attachment security for cognitive and language outcomes of low-income toddlers." *Infant Behavior and Development*, 26, 326-344.

Spiess, C. and K. Wrohlich (2008). "The parental leave benefit reform in Germany: costs and labour market outcomes of moving towards the Nordic model." *Population Research and Policy Review*, 27, 575-591.

Vandell, D., Burchinal, M., Friedman, S., and C. Brownell (2001). "Overview of early child care effects at 4.5 years." In meeting of the Society for Research in Child Development, Minneapolis, MN.

Table 1: Sample means of time-constant controls

	Was enrolled in infant daycare n = 560	Was not enrolled in infant daycare n = 1,370
Age in 2008	25.9	26.0
Female	0.49	0.49
Black	0.19*	0.12
Hispanic	0.09*	0.14

* First column differs from second column at $p < .05$

Table 2: Sample means of time-varying controls

	Was enrolled in infant daycare N = 3,920	Was not enrolled in infant daycare N = 9,590
Married	0.48	0.47
Any kids	1.06*	1.13
Regional unemployment rate	7.23	7.23
Regional percent change in housing starts	−4.31	−4.41

* First column differs from second column at $p < .05$

Table 3: Sample means of outcome variables

	Was enrolled in infant daycare N = 3,920	Was not enrolled in infant daycare N = 9,590
Employed full time	0.67*	0.63
Log income (2017 dollars)	10.79*	10.71
Above 200 percent of poverty line	0.73*	0.70

* First column differs from second column at $p < .05$

Table 4: Intra-person coefficients of variation for time-varying variables (intra-person standard deviation divided by overall mean)

Employed full time	0.52
Log income	0.11
Above 200 percent of poverty line	0.43
Married	0.55
Any kids	0.57
Regional unemployment rate	0.28
Regional percent change in housing starts	4.74

Table 5: Estimates of the effects of first-year daycare enrollment on later-in-life outcomes

	Full time		Log income		Above 200% poverty	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Constant only	0.039**	0.017	0.085	0.054	0.037**	0.016
Controls	0.037**	0.017	0.106**	0.048	0.034**	0.014
Controls + lagged outcome	0.015*	0.008	0.052*	0.030	0.019**	0.008
Controls + lagged outcome + fixed effects	−0.080	0.186	0.090	0.683	0.077	0.209

* $p < .10$

** $p < .05$

Appendix Table 1: Estimation results
Outcome: employed fulltime?

	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Time-varying						
Lagged fulltime					0.580	0.010
Married			0.099	0.015	0.042	0.008
Number of kids			−0.039	0.006	−0.016	0.003
Regional percent change in housing starts			0.014	0.008	0.017	0.007
Regional unemployment rate			−0.005	0.002	−0.0001	0.002
Time-constant						
Age			0.010	0.006	0.0004	0.003
Female			−0.177	0.015	−0.083	0.008
Black			−0.003	0.022	0.002	0.011
Hispanic			0.017	0.023	−0.001	0.011
First-year daycare enrollee	0.039	0.017	0.037	0.017	0.015	0.008
Constant	0.627	0.009	0.487	0.143	0.281	0.068

Appendix Table 2: Estimation results
Outcome: log income

	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Time-varying						
Lagged fulltime					0.463	0.026
Married			0.745	0.044	0.479	0.033
Number of kids			−0.083	0.019	−0.053	0.013
Regional percent change in housing starts			0.025	0.030	−0.022	0.027
Regional unemployment rate			−0.019	0.007	−0.014	0.006
Time-constant						
Age			0.039	0.017	0.018	0.010
Female			0.029	0.045	0.018	0.028
Black			−0.667	0.096	−0.391	0.059
Hispanic			−0.327	0.083	−0.202	0.052
First-year daycare enrollee	0.084	0.054	0.106	0.048	0.052	0.030
Constant	10.706	0.031	9.665	0.443	5.338	0.398

Appendix Table 3: Estimation results
Outcome: Above 200 percent of the poverty line

	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Time-varying						
Lagged fulltime					0.474	0.012
Married			0.240	0.012	0.134	0.008
Number of kids			−0.114	0.005	−0.067	0.003
Regional percent change in housing starts			0.025	0.007	−0.005	0.007
Regional unemployment rate			−0.011	0.002	−0.009	0.001
Time-constant						
Age			0.018	0.005	0.008	0.003
Female			−0.011	0.013	−0.008	0.007
Black			−0.142	0.022	−0.082	0.013
Hispanic			−0.075	0.021	−0.050	0.012
First-year daycare enrollee	0.037	0.016	0.034	0.014	0.019	0.008
Constant	0.698	0.009	0.348	0.125	0.279	0.072

Appendix Table 4: Estimation results
Two-step dynamic fixed effects estimates

	Full time		Log income		Above 200% poverty	
	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
First stage (time-varying)						
Lagged outcome	0.404	0.018	0.274	0.040	0.216	0.016
Married	0.020	0.014	0.261	0.038	0.065	0.014
Number of kids	−0.013	0.006	0.087	0.028	−0.063	0.006
Regional percent change in housing starts	0.017	0.008	−0.011	0.027	0.002	0.209
Regional unemployment rate	−0.001	0.002	−0.015	0.008	−0.013	0.002
Second stage (time-constant)						
Age	0.002	0.004	0.023	0.017	0.015	0.005
Female	−0.117	0.011	−0.033	0.042	−0.018	0.012
Black	−0.001	0.024	−0.636	0.105	−0.165	0.029
Hispanic	−0.009	0.020	−0.321	0.084	−0.078	0.023
Daycare	−0.080	0.186	0.090	0.683	0.077	0.209
Constant	0.408	0.136	7.245	0.651	0.323	0.159