Dynamic Spatial Competition in Early Education: an Equilibrium Analysis of the Preschool Market in Pennsylvania*

Pierre Bodéré†

22nd November, 2023

Updated regularly, please click here for the latest version.

Abstract

High-quality preschool is one of the most cost-effective educational interventions, yet the United States invests little in early childhood education. Recent policy discussions call for increasing preschool enrollment and raising the quality provided, especially for disadvantaged children, but equilibrium responses of private providers which make up most of the market generate trade-offs between these objectives. Supply expansion may lower incentives to invest in quality, and price responses to demand subsidies can increase the costs faced by non-subsidized parents. This paper develops a dynamic model of the preschool market to evaluate the effectiveness of policies at achieving these objectives. The model nests a static equilibrium model of spatial competition and preschool choice within a dynamic model of providers' entry, exit and quality investments. I estimate this model using data on the universe of child-care centers in Pennsylvania. I use the model to simulate the aggregate and distributional consequences of proposed approaches to early education expansion. I find that policies focused on expanding supply raise access but decrease the quality children are enrolled in due to parents' value for proximity. Demand subsidies generate market expansion, but on their own do not create sufficient incentives for providers to invest in quality. Among the simulated policies, the most cost-effective at expanding high-quality enrollment combine demand subsidies targeted to low-income families with financial support to high-quality providers serving disadvantaged children. These policies increase access by reducing exit of providers, and expand high-quality enrollment for low-income children through subsidies. In addition, these targeted policies generate spillovers to the educational quality of non-targeted families by creating incentives for centers to invest in quality.

*I am grateful to Chris Conlon, Guillaume R. Fréchette and Daniel Waldinger for the constant guidance and support. This project greatly benefited from numerous conversations with Claudia Allende, Milena Almagro, Tslil Aloni, Elizabeth Davis, Mike Dickstein, Chris Flinn, Alejandro Ganimian , Mike Gilraine, Elena Manresa, Marguerite Obolensky, Aaron Sojourner, Paul Scott, and Teresa Steininger, and from audiences at NYU, Stern and at the Kumar Conference on Early Childhood Development. I am especially indebted to Jonathan Elliott for his invaluable help with gathering the data used in this paper.

†Department of Economics, New York University. E-mail: pmb412@nyu.edu

1 Introduction

High-quality early childhood education (ECE) provides benefits that last long into adulthood, in particular for disadvantaged children (Elango et al., 2015). The successes of high-quality preschool hold the promise of reducing educational inequality at little expense: these interventions are among the most cost-effective social programs (Hendren and Sprung-Keyser, 2020). Beyond low-income families, the availability of high-quality ECE constitutes a crucial substitute for parental time now that the majority of households are dual earners. Despite these excellent social returns, the U.S government spends relatively little on early education: $\$2,800$ a year per child aged 3 and 4 compared with $\$12,800$ on elementary-aged children.[1] These low levels of public spending result in an early education market characterized by the importance of private provision, low enrollment, low quality, and high prices for families.[2] Given the many barriers to high-quality early education young parents face, calls for increased investments in preschool have emerged from policy makers and scholars alike (Department of the Treasury, 2021; Davis and Sojourner, 2021).

In early education, private provision plays a central role and profit margins are thin.[3] Policies can thus result both in short-run equilibrium responses and long-run changes in the characteristics of the preschool supply. Providers with market power may raise their prices, which can undo some of the benefits of policies. Providers are also likely to enter more, exit less, and invest more in quality in areas made more profitable either by parents' demand or by public funds. The impact and distributional consequences of policies thus depend on the costs providers face to open, operate, and raise the quality of a preschool, and on how parents from different demographics decide among early education arrangements.

This paper develops a dynamic spatial competition model to study the equilibrium implications of early education policies. This framework endogenizes three key determinants of parents' early education choices: affordability, access, and quality. I estimate the model using comprehensive data on the universe of ECE providers in Pennsylvania. With the estimated model, I run counterfactual simulations to evaluate the impact of proposed approaches to early education expansion on aggregate preschool enrollment and quality, and on their distribution across family income. The results reveal trade-offs between the policy goals of expanding access and enrollment for all, improving early

[1] See Figure 2 in Davis and Sojourner (2021)

[2] See section 2.1 for a discussion on private provision and enrollment rates. According to the National Center for Education Statistics, less than a third of center-based arrangements are rated high quality (U.S. Department of Education, 2014).

[3] Grunewald and Davis (2011) review a series of facts on early education providers' profits and turnover.

education quality, and targeting low-income children. Entry subsidies increase access and parents' welfare as the entry of providers create new options nearby, but also result in low-quality entrants stealing business from higher quality centers. Demand subsidies are necessary to raise enrollment, but do not constitute a cost-effective way to improve the quality supplied because of providers' price responses. In contrast, policies that combine targeted demand subsidies with financial incentives for high-quality centers serving low-income children generate incentives for providers to invest without distorting prices, thus raising high-quality enrollment for all income groups.

Pennsylvania has several attractive features for this purpose. The state combines a quality rating system (QRIS) which monitors and communicates providers' quality to parents, a nexus of demand and supply side policies providing sources of variation of centers' revenues, and rich administrative data on centers' prices and enrollment by children's income. Descriptive analysis of this data reveals three main findings. First, geographic access to ECE, including to high-quality centers, is unequally distributed. Neighborhoods with more college educated women have more centers within a 10-minute drive, in particular high-quality ones. This result is robust to measures of access adjusting local supply by the local population of preschoolers. Second, these disparities result from differential quality upgrade and exit decisions by providers across areas, with centers exiting more and upgrading less in lower income neighborhoods. Lastly, providers' upgrade choices respond to financial incentives induced by policies, with an extra $100 in daily subsidies translating into a 2.1 p.p increase in the rate of upgrade to the highest rating.

Motivated by this evidence, I develop a model of supply and demand of preschool education. The model comprises two nested parts. First, a static model of competition endogenizes equilibrium prices and enrollment, taking the available supply of early education as given. This first modeling step delivers the variable profits earned by providers in any state of the preschool market. Second, a dynamic model of entry, exit, and quality investments determines the evolution of the distribution of centers and their quality. Providers make these decisions by comparing the expected profits from an action to its corresponding cost.

The static component of the model captures parents' choices and providers' pricing decisions in a differentiated product competition framework. On the demand side, households differing by human capital, income, and location decide whether and to which center to send their child. On the supply side, providers are differentiated across space, observed rating and unobserved quality, and compete on prices to attract children. The model incorporates policies on both sides of the market: demand subsidies which reduce consumer prices for low-income parents, and financial support to providers which raises producer prices of high-quality centers. This allows me to disentangle the role of policies from that of preferences and costs in determining enrollment and prices, and to evaluate how these equilibrium outcomes would respond to an alternative environment.

I endogenize the ECE market structure with a dynamic discrete game played among providers. Every year, short-lived potential entrants decide whether to enter a specific neighborhood. Simultaneously incumbents choose whether to remain in the market, and if so the quality at which to operate next year. Centers make these decisions strategically, taking into account the impact of nearby competitors on expected future profits. The policy environment might influence these decisions both through static and dynamic channels. An entry or investment policy (dynamic) can directly influence market structure by reducing the costs of entering or raising quality. An expansion of subsidies to disadvantaged families (static) may increase enrollment in existing centers, but the take-up could be limited due to parents lacking nearby options. However, this policy also changes incentives for centers to operate, and the entry of new providers may increase enrollment over the years. The dynamic responses of providers can thus amplify the benefits of early education policies directed at static incentives.

Computing counterfactual market structure under new policy environments requires solving this dynamic game, which poses a methodological challenge. The state space is high-dimensional, as the information relevant for preschools' decisions includes a large number of discrete and continuous variables such as characteristics of all competing centers in the market. I combine parametric policy iteration (PPI) with value function approximation to make the solving the game computationally tractable (Rust, 2000; Sweeting, 2013). I train neural networks to compute a fast and reliable approximation of variable profits for any given policy environment and state of the market. I then use the trained network to construct relevant bases of approximation for the value function of preschool operators. This approach makes the computation of centers' equilibrium strategies feasible while preserving the richness of the static and dynamic competition between providers.

Estimation proceeds in two-steps. First, the static parameters governing preschools' marginal costs and parents' preferences are recovered jointly using a Generalized Method of Moments estimator. I use variation over time and across space in cost-shifters such as teachers' wages, and in local competition to form instruments and address endogeneity of providers' prices and enrollment. The estimated preference parameters highlight two important dimensions of heterogeneity in the way parents trade-off attributes of ECE providers. Non-college educated parents value proximity more, while college-educated parents place a greater value on highly rated centers. Estimates of marginal costs suggest that providing early education at the highest rating of the Pennsylvania system costs \$6 more on average than at the lowest rating. Given estimates, the static model allows me to compute the variable profits of centers for any given market state, a prerequisite for the estimation of the dynamic parameters.

In a second step, I estimate the dynamic parameters consisting of the entry costs, centers' fixed operating costs, upgrade costs to higher ratings, and the variance of private payoff shocks. I follow Sweeting (2013) and use a combination of PPI with a variant of the Nested Pseudo-Likelihood (NPL) algorithm of Aguirregabiria and Mira (2007, 2010) in a full solution approach. This procedure jointly searches for parameters and for centers' best responses. Each step, the approximation to the value function is updated and new parameters are computed by maximum likelihood. The procedure yields parameters and strategies consistent with both equilibrium restrictions and the transitions observed in the data. Combined estimates of fixed and marginal costs suggest an annual cost of ECE provision of around $\$8,700$ per child, close to industry estimates for Pennsylvania.[4]

Policy Implications. The model endogenizes three outcomes of particular importance for the design of early education policies: affordability, access, and quality. I exploit this richness to evaluate two policies working through static channels (expansion of subsidies to middle-income families, and targeted financial support to providers) and one policy directed at preschools dynamic decision (start-up grants). The design of these counterfactuals is inspired by policies that are either discussed or implemented in the United States.[5] To evaluate each environment, I compare it to a baseline scenario where parents' preferences and preschools' costs are the only determinants of equilibrium outcomes. I form several measures of the cost-effectiveness of these policies: dollars spent per additional preschooler, per preschooler in high-quality, and dollars spent per new preschool. I also use the demand model to look at the distributional consequences of these policies on (high-quality) enrollment by family income.

Two of the proposed policies, start-up grants and demand expansion to middle-income families, represent large investments in early education. These policies are not cost-effective at increasing enrollment, but their benefits and shortcomings shed light on the forces shaping equilibrium outcomes in early education markets. Start-up grants generate entry of a large number of new preschools at

[4] Friedman-Krauss et al. (2018)

[5] The expansion of subsidies to middle-income families transposes to preK the stated goal of Biden administration for childcare: "*to provide access to high-quality child care for low- and middle-income children*" (Department of the Treasury, 2021). Targeted financial support to providers is similar to the incentives used in the majority of QRIS, the main market-based approach to early education policy in the U.S. Start-up grants have been implemented in a variety of settings such as Florida, Indiana, Pennsylvania, and more recently New York (Tout et al., 2010).

a relatively low-cost per center. This raises parents' welfare and increases enrollment, in particular for low-income families, as parents value having convenient options nearby. However, the increased competition from new entrants displaces enrollment from high- to low-quality centers. Expanding subsidies to middle-income parents creates market expansion and welfare gains for this group of parents, but makes early education less affordable to high-income parents due to providers' prices responses. These price adjustments are responsible for the high-cost of the policy. Taking into account providers' dynamic responses reveals an additional long-term effect of the policy which mitigates this equity-efficiency trade-off. By increasing the demand from quality sensitive middle income families, this policy generates incentives for centers to invest, which raises the quality attended by low-, middle-, and high-income children.

The policy that performs the best in terms of cost-effectiveness, in particular for high-quality enrollment, is the targeted financial support to providers combined with subsidies directed at low-income households. Expanding on the tiered-reimbursement scheme in place in Pennsylvania called add-on rates, this policy provides a substantial amount of additional funding to centers serving low-income children at high-quality. Subsidies to low-income parents are required to alleviate the costs faced by these families, bring their children in the market, and make add-on rates a source of revenue for centers. By tying public funds to both high-quality provision and low-income enrollment, this policy breaks the spatial matching between families' socioeconomic status and centers' revenue gains from investing in quality. The policy is also effective at raising high-quality enrollment, and does so not only for targeted children, but also for non-subsidized families. First, the incentive scheme on the supply side acts as a substitute to higher markups and therefore mitigates the price responses from providers, lowering the costs for non-subsidized families. Second, neighborhoods are not completely segregated by income, implying that the higher upgrade rates of centers having part of their enrollment low-income results in more high-quality centers available for middle- and high-income parents. These results are encouraging in light of the deployment of QRIS in most U.S states in the past few years.[6]

[6] See Figure G.1 in Appendix for the deployment of QRIS in the U.S.

Related Literature. This paper builds on the findings from an expanding literature studying the impact of early childhood education programs on child development in North America. Seminal studies evaluating the impact of high-quality, targeted, small-scale early intervention programs such as Perry preschool or the Infant Health and Development Program find large benefits for participants on a battery of short and long-run outcomes (Heckman et al., 2013; Duncan and Sojourner, 2013; Elango et al., 2015). Head Start, the federal early childhood education program targeted to low-income children similarly delivers large gains for enrolled children in the short and long-run (Kline and Walters, 2016; Ludwig and Miller, 2007). Evaluations of large-scale, universal programs implemented in Georgia, Oklahoma, and Boston lead to large benefits for children while programs in Tennessee or the province of Québec proved detrimental to participants (Cascio and Schanzenbach, 2013; Gray-Lobe et al., 2021; Durkin et al., 2022; Baker et al., 2008, 2019). Cascio (2015) interprets these findings as follows: preschool contributes to child development only when provided at high-quality, and the gains depend on the quality of available alternatives. Building on these results, this paper endogenizes the main outcomes of interest for policy makers pursuing the goal of improving child development through preschool expansion: the quality provided by centers, and the demographics of children served.

This paper contributes to the growing literature studying the supply side of early childhood education. Several studies highlight the importance of considering providers' dynamic responses to a change in the policy environment. Bassok et al. (2014) find that competition with the public sector can crowd-out private preschools. Brown (2018) shows that a reduction of centers' revenue can lead to lower quality, reduced capacity or exit altogether. Hotz and Xiao (2011) show that more stringent quality regulations lead to a decrease in the supply of ECE in disadvantaged neighborhoods but an increase in quality in affluent areas. This paper builds on these findings and incorporates providers' dynamic adjustments in the design of ECE investments. A recent group of papers studies the effects of ECE expansion using general equilibrium models including labor force participation (Borowsky et al., 2022; Berlinski et al., 2020). In contrast, this paper focuses on the supply-side responses to ECE investments. In addition to parents' preschool choices, I endogenize key margins of differentiation between providers: entry, exit and quality investments which together determine the spatial distribution of access, and prices, which result from competition between providers with market power.

This paper extends the empirical industrial organization literature studying education markets in equilibrium by incorporating providers' dynamic decisions. Neilson (2013); Allende (2019) study schools responses to policies using a differentiated product model of price and quality competition between schools. Armona and Cao (2022) study the design of federal aid in a context in which for-profit colleges set both tuitions and advertising levels. More closely related to this paper, Singleton (2019) and Dinerstein et al. (2020) combine school choice frameworks with models of entry and exit of education providers. In both settings, these decisions are modeled with a static game. In contrast, I model the dynamic considerations of early education providers when making their decision to enter, to exit, or to switch the quality they provide. This allows me to study the impact of ECE policies taking into account both short-run price and enrollment responses, and long-run changes in the supply of quality across space. Moreover, the dynamic framework makes it possible to compare the costs of policies operating through static channels, such as demand subsidies, which are paid every year, and dynamic policies, such as entry or investment subsidies, which are paid only once. The combination these two features makes this framework well-suited to evaluate the large range of policy instruments used in ECE.

Finally, this paper contributes to the empirical literature on dynamic games. One of the main challenges in this literature lies in the tractability of the game's solution. Most applications avoid the costly computation of value functions and resort instead to two-steps estimation methods in which first-stage conditional choice probabilities (CCP) are mapped to value functions. This approach is used to analyze the environmental regulation of cement by Ryan (2012), to study of demand fluctuations in ready-mix concrete by Collard-Wexler (2013), or spatial competition in groceries by Caoui et al. (2022). These papers rely on a reduced-form model of profits. Investments in preschool such as demand subsidies may influence competition between providers, impacting parents' choices, prices charged, and profits. This dependence creates the need for a structural model of competition between firms. Policies can also directly or indirectly influence providers' dynamic decisions, making first-stage CCPs poorly suited to model firms' actions in a counterfactual scenario. This point causes me to adopt a full-solution approach to simulate and estimate the model. I follow Sweeting (2013) and rely on a combination of value function approximation and NPL. To speed up the computation of profits, which depend on the equilibrium of a Nash pricing game, I use neural networks to provide a fast and reliable approximation of profits in any state. The combination of the external validity of equilibrium models with the speed and predictive performance of neural networks opens up possibilities to simulate counterfactual environments in games with large state spaces.

2 Institutional Context and Data

2.1 ECE Landscape in the U.S. and Policy Objectives

Following the increase in female labor force participation during the second half of the XXth century, the majority of children in high-income countries spend several days a week in non-parental care (Blau and Currie, 2006). In the United States, a variety of arrangements have developed for children from birth through age 5, ranging from informal options such as care by a relative to more organized programs. This paper focuses on preschool, a specific arrangement which the U.S Census Bureau defines as "*a group or class that is organized to provide educational experiences for children during the year or years preceding kindergarten which includes instruction as an important and integral phase of its program of child care*".[7] This definition excludes home and family-based services. The focus on preschool stems from two reasons. First, formal environments tend to be of higher quality than informal arrangements (Bernal and Keane, 2011), thus providing a better policy instrument for improving child development. Second, center-based care corresponds to a well-defined industry with a common production structure, which facilitates the estimation of costs. Throughout this paper I use the words preschools, pre-kindergarten, and centers interchangeably.[8]

In contrast with kindergarten, preschool is far from commonplace as an educational experience for American children. Kindergarten rapidly expanded during the 1960s and is now an integral part of the U.S public education system, often referred to as K-12, with all states funding a form of kindergarten (Kamerman and Gatenio-Gabel, 2007). As a result, more than 90% of 5 year olds are enrolled in school (National Center for Education Statistics, 2022). Achieving a similar momentum has proved more difficult for early education before age 5, in part due to historical reasons. ECE in the U.S. originates from two movements with differing purposes: nursery schools and day nurseries. Nursery schools focused on children's educational experiences and were increasingly adopted by middle and upper-class families. Custodial in nature, day nurseries were developed to cater full-time for children from low-income families with working mothers. Many of these programs did not provide an adequate environment for child development. Day nurseries later gradually integrated the educational practices from nursery schools. But the adverse outcomes observed for the children placed in the low-quality institutions created a stigma associated with public ECE, and called into

[7] US Census documentation

[8] The distinciton between centers and preschool is described as "at best fuzzy"(Kamerman and Gatenio-Gabel, 2007). In addition, it does not correspond to any observable categories in the data used in this paper and is therefore ignored.

question non-parental care at such early ages (Cahan, 1989). Debates over the high cost of publicly funded preschool and the role the government should play in early education also contributed to the difficulty to secure federal support (Beatty, 1995).[9] These developments have resulted in a fragmented educational landscape before kindergarten.

Compared to K-12 education where the vast majority of children attend public schools, market provision is central to the ECE system. More than 80% of child care centers are run by private entities (NSECE Project Team, 2014).[10] In addition to being privately provided, ECE is the only segment of the education pathway for which families bear the majority of the costs. An estimated 52% of spending in ECE comes from private payments (BUILD Initiative, 2017), and most centers rely on tuition as one of their main sources of revenue (NSECE Project Team, 2014). Mandatory staff-to-children ratios which are parts of licensing requirements contribute to making ECE expensive to provide. As a result, high costs are reported as the main barrier to early education by families, with long distance and low quality ranking second and third (Cui and Hanson, 2021). These market conditions result in low preschool enrollment of 3 and 4 years: the U.S ranks 41st and 43rd among OECD countires in enrollment rates of these age groups. (OECD, 2019). These attendance rates are stable since the early 2000s, averaging respectively around 40% and 65%.[11] Preschool attendance correlates with household demographics. Children in more educated households are more likely to be enrolled, be it half-day or full-day (McFarland et al., 2019). Figure 1 shows the increase in the share of both 3 years old and 4 years old attending pre-kindergarten with household income. However, not even among the richest U.S households does the share of enrolled preschoolers reach the OECD average.

In spite of the prevalence of private provision and reliance on payments from families, public funding plays an important role. The two main federal programs of funding for ECE are Head Start and, more relevant to this paper, the Child Care Development Fund (CCDF).[12] Head Start consists in a combination of education, nutrition, health and parent involvement services to low-income households. Head Start has restrictive eligibility requirements and serves a million children, roughly 5% of Americans under age five. The program provides high-quality care and is associated with improved outcomes for participants. As such, Head Start can be seen as a direct provision

[9]In 1971 President Nixon vetoed the Comprehensive Child Development Act on the basis that "Good public policy requires that we enhance rather than diminish both parental authority and parental involvement with children." (Rosenthal, 1971)

[10]These non-governmental entities include independent for-profit centers, franchises of large chains, and non-profit centers which can be part of a broader faith or community-based organization.

[11]NCES Preschool Enrollment Summary

[12]Hotz and Wiswall (2019) provide an illuminating review of ECE policies in the U.S.

policy simultaneously addressing cost, distance and quality for low-income families. The CCDF provides grants to states to subsidize ECE for low-income parents. States have discretion on the eligibility requirements, but most recipients served are below 150% of the federal poverty line (FPL). States also distribute grants to providers. The goal of these two components of the CCDF are to decrease the costs to families through subsidies, and to stabilize the supply of ECE through financial support to providers. A last aspect of the CCDF of particular interest for this paper is the funding of states' quality rating and improvement systems (QRIS). QRIS are market-oriented systems which seek to improve the quality of ECE. These systems involve distributing funding and implementing incentive schemes for providers in addition to communicating information about education quality to parents. All U.S. states except Mississippi have a QRIS either planned or implemented (Cannon et al., 2017). Direct provision and market-based QRIS both pursue the objective of providing high-quality early childhood education to low-income parents.

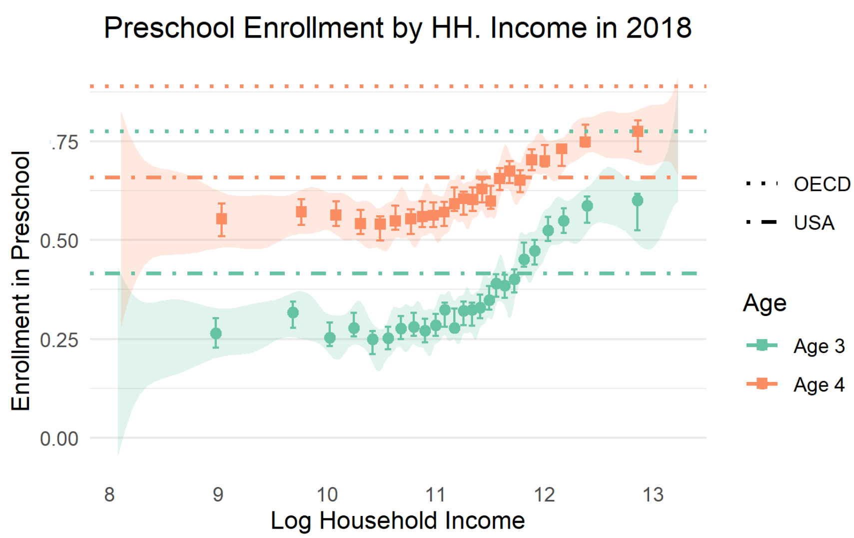


Figure 1: Share of Children Enrolled in Preschool by Age, U.S 2018

2.2 The ECE Market in Pennsylvania

This paper focuses on the center-based ECE market in Pennsylvania. The state implements a QRIS which was launched in 2002. The main demand subsidy branch of this QRIS is an income-based subsidy to families called Child Care Works (CCW).[13] This policy reduces the consumer price faced by subsidized parents relative to the market price which is paid in full by richer families. The amount

[13] The state also implements a direct provision system called PA-PreK Counts, but the enrollment in that program is low (Friedman-Krauss et al., 2018). The paper this ignores this branch of the policy to focus on CCW.

paid by subsidized parents is the maximum of a co-payment and the market price from which is subtracted the maximum child care allowance (MCCA), a subisdy ceiling which varies across regions. Under this policy, a hypothetical family with a yearly income of $\$33,000$, corresponding to a $10 daily co-copayment, in a region with a MCCA of $32, sending her child to a center with a market price of $35 would end up paying for $\max\{copay, p - MCCA\} = \max\{10, 35 - 32\} = \$10$.

The quality rating system in Pennsylvania is called Keystone Stars. The system rates centers on a four-levels scale, where STAR 1 is the lowest possible rating and STAR 4 the highest.[14] The quality ratings encompass 4 dimensions of improvement for providers: (1) staff qualifications, (2) early care education program, (3) partnership with families, and (4) management practices. At STAR 3 and 4, the first dimension requires that a specified fraction of employees must have completed or be enrolled in a child development program. The education program component requires centers to periodically have an assessor score the quality of teaching on measures such as the environment rating scale (ERS). ERS is part of the metrics called *process measures of quality* by the early learning literature. Process measures positively impact child development, although effect sizes tend to be small (Keys et al., 2013).

Programs can apply for a higher rating any time. Upgrading quality incurs various sunk costs, ranging from preparing the certification requirements to raising staff qualifications. Higher quality is also likely associated with higher operating costs stemming from increased wages and more stringent standards. A center can downgrade to a lower rating if the requirements to maintain its current rating are not met, even though downgrading is rare in the data.

Keystone Stars is supplemented by policies creating incentives for providers to operate at high-quality, in particular when serving low-income children. This incentive scheme takes the form of quality specific add-on rates. Figure B.1 shows the schedule of add-on rates over the years covered in the data. I assume that centers always receive the maximum add-on rate associated with their rating, regardless of their market price, and that add-ons do not influence family payments.[15] With this assumption, in 2018 a STAR 4 provider with a market price of $45 serving the same hypothetical family as in the previous example would receive $13 + 32 + 9.2 = \$54.20$ where the first term corresponds to the family payments, and the second and third terms, both paid by the state, correspond to the subsidy and the add-on rate. In contrast, the same center would only receive $45

[14] No STAR, a possible category before 2016, disappeared to be combined with STAR 1. To facilitate the analysis, I bundle No STAR and STAR 1 since the start of the data.

[15] In practice, QRIS tend to cap payments to providers so that amounts received do not exceed market rates. My working assumption abstracts from this, but still captures the intended goal of add-on rates, and makes demand and supply side subsidies additive and independent. I will incorporate this constraint in a future iteration of the paper.

from serving a non-subsidized child, and family payments would cover the whole amount.

2.3 Data Sources and Descriptive Statistics

This paper relies on data sources covering the supply and demand sides of the ECE market in Pennsylvania. This section describes the data used Table 1 provides summary statistics of the main variables of interest.

Pennsylvania QRIS Data. The main dataset used in this paper covers the main elements of the Keystone Stars system described in the previous paragraph. This includes detailed information on providers over the years 2010-2018, and on the time series of policies implemented over that period. The data were obtained through Freedom of Information Act request to the Office of Child Development and Early Learning of Pennsylvania. The provider level variables are observed on a yearly basis and include price charged for each age group, enrollment broken down by age and subsidy status of children, and STAR rating. Enrollment is collected at the classroom levels within provider, and contains the number of teacher and aides in each classroom. The enrollment is measured in full-time equivalent units. The join observation of price, enrollment, and a qaulity rating measure is a rare feature in early education. But a limitation of this data is that STAR 1 providers, which are not eligible for supply side incentives, do not have to report their enrollment. I detail how the estimation approach used in the paper addresses this fact in Appendix E.2. In addition, I observe licensing information such as the capacity, ownership structure and for-profit status of the provider. The policy variables of interest include for each year the income-based copayment schedule determining subsidized families payments in each year, the maximum child care allowance, and the schedule of quality add-on rates. The sample is restricted to providers of the center, excluding home and family based ECE.

Market Demographics. I use the American Community Survey (ACS) to get demographic information on the demand-side of the preschool market. The main analysis in the paper relies on a spatial demand model, which involves simulating preschoolers at geographically disaggregated level. I use 5-year ACS block group estimates, imputed with tract-level variables when the block group information is not available.[16] The main demographic variables of interest are population by age, gender, income, and educational attainment.

[16] See Appendix ? for the imputation procedure used for education-contingent income distributions at the block-group level.

Additional Data Sources. The National Survey of Early Care and Education (NSECE) is used to construct additional moments on the characteristics of ECE arrangements by demographics. This survey, respresentative at the national level, contains detailed information about young parents ECE choices, including distance traveled and costs. This information supplements the data presented above, which is detailed at the provider level but does not contain individual choices.

Descriptive Statistics. There are almost 4000 centers serving a population of 3 and 4 year old children in the last year of the sample. The quality rating and improvement system (QRIS) in place in Pennsylvania applies a four level grading system called Keystone Stars to licensed early education providers, with STAR 1 designating the lowest level and STAR 4 the highest. Panel A of Table 1 shows the characteristics of prekindergartens operating in 2018 by Keystone Stars rating. The majority (2028) of centers operate at the lowest level of quality, followed by the highest quality tier (751). Characteristics of providers differ across quality ratings. Average full-time daily prices for preschool-aged children range from \$36.7 for STAR 1 providers to \$42.7 for the highest rating. Highly rated centers tend to be larger structures: the average licensed capacity of STAR 4 centers is 124.3 while that of STAR 1 centers is 70.9. Finally, more than a third of STAR 4 prekindergartens are accredited by the National Association for the Education of Young Children (NAEYC), a measure of quality used across the country. Panel B of Table 1 presents statistics constructed from classroom level yearly reports submitted by providers rated STAR 2 and above. Patterns of enrollment coincide with licensed capacity measures: highly rated centers tend to serve more preschoolers, have a larger teaching staff and also more classrooms dedicated to preschool. While licensed capacity, which relates to the building's characteristics, tends to be high, the potentially binding constraints limiting a center's enrollment are state mandatory staff-to-children ratios and group sizes. For preschoolers, the maximum staff to children ratio is 0.1 and group size is 20, implying that a provider enrolling 40 preschoolers must have at least 2 classrooms dedicated to this age group run by 2 teachers each. Across quality ratings, only 20% of providers operate at 90% of their maximum capacity.

Table 1: Descriptive Statistics of Prekindergartens in 2018

	STAR 1 Centers	STAR 2 Centers	STAR 3 Centers	STAR 4 Centers
		Panel A: Center Characteristics		
Number of Centers	2028	690	459	751
Preschool Daily Price	36.7	35.87	39.69	42.74
Licensed Capacity	70.9	85.06	113.08	124.32
Accreditation Status	0.03	0.02	0.01	0.35
		Panel B: Enrollment Characteristics		
Full Time Eq. Preschool Enrollment		22.69	35.71	43.01
Share Private Pay Enrollment		0.51	0.51	0.52
Number Preschool Teachers		3.53	5.25	6.01
Number of Preschool Classrooms		1.98	2.87	3.25
Share 90% of Max Capacity		0.19	0.17	0.2

Notes: Enrollment characteristics are computed from classroom level reports sent by Keystone Stars providers. STAR 1 providers are not compelled to such reports are thus are excluded from this table. Max capacity is determined from staff-to-children ratios and not licensed capacity, which is seldom binding.

3 Stylized Facts

This section provides descriptive evidence about the preschool market in Pennsylvania. The first fact is that access to preschool is unequally distributed across space, with more educated neighborhoods having access to more centers, and more high-quality centers in particular. Second, the quality rating used in Pennsylvania appears to be valued by parents, especially those with college degree. Third, providers respond to a change in the policy environment raising incentives to operate at higher quality tiers by investing in quality and increasing prices. Together, these facts provide guidance for the design of the industry model presented in Section 4.

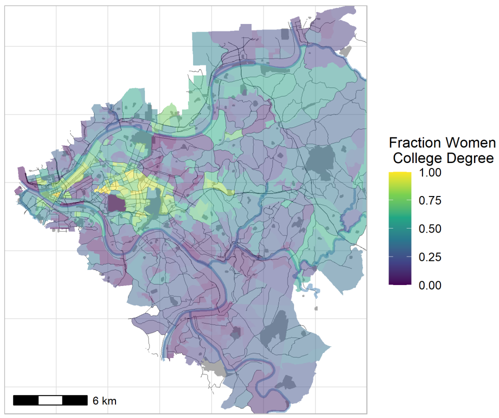
3.1 Access to Preschool is Unequal

Physical distance to centers plays a key role in shaping parents' choices of ECE provider. Responses to the National Survey of Early Care and Education (NSECE) from 2012 indicate that 55 percent of early education arrangements for children age 3 through 5 years are located within 3 miles of the child's home (OPRE, 2016).[17] The importance of distance has spurred an extensive literature studying disparities in supply of ECE services available to parents of different socio-economic status (Bassok et al., 2011; Bassok and Galdo, 2016; Davis et al., 2019).

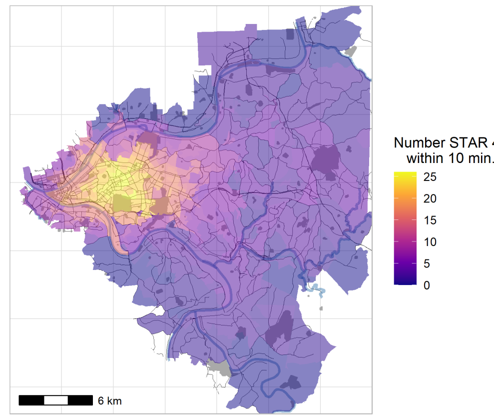
Figure 2 illustrates how access inequality materializes in Pittsburgh in 2018. Maps (a), (b) and (c) focus on the eastern part of the city delimited by the Allegheny river to the north, the Monongahela river to the south, and by their confluence to the west. The top-left panel displays the distribution of college educated women by census block-groups. The majority of college graduates reside in the area neighboring the campuses of Carnegie Mellon and of the University of Pittsburgh. Panel (b) shows the number of STAR 4 preschools located within a 10 minutes drive from the census block-group centroid. Neighborhoods where high-quality centers are concentrated are also the areas with a high-fraction of college educated women identified in panel (a). In contrast, most census block-groups in the southern and eastern peripheries of the market have limited high-quality options nearby. This pattern presented for STAR 4 centers also holds true for the total number of preschools in general.

To argue that this correlation is not merely due to more preschoolers residing in highly educated neighborhoods, I borrow from Davis et al. (2019) a measure of demand-adjusted supply called access. For each preschool j of capacity C_j, and for each census block-group ℓ with a population

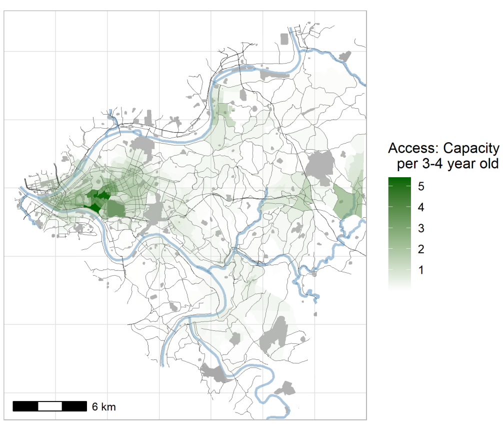
[17] NSECE is a survey designed to document patterns of use and availability of care in the United States. The survey consists in 3 components: Household, Provider, Workforce.



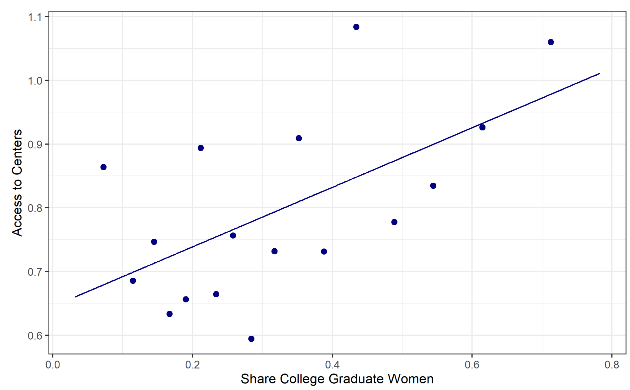
(a) Share of College Educated Women



(b) STAR 4 Centers Supply



(c) Access



(d) Access and Education, Pennsylvania

Figure 2: Heterogeneous Access to Early Education

Notes: Top left panel: Share of college educated women in census block-groups in Pittsburgh. Top right panel: Number of STAR 4 centers within a 10 minutes travel time from the census block-group centroid. Travel time is computed using Open Street Map. Bottom left panel: Access is defined following the Enhanced Two Step Floating Catchment Area method. Bottom right panel : Binscatter plot of access to preschool as a function of the share of college educated women in a census tract, for all of Pennsylvania, 2018. The number of bins is chosen using the Cattaneo et al. (2019) data-driven approach.

N_ℓ of preschoolers I define slot per capita at the school (S_j) and block group (A_ℓ) levels as follows:

$$S_j = \frac{C_j}{\sum_{\ell,\, t_{\ell j}<20 \text{ min}} w(t_{\ell j}) N_\ell} \text{ and } A_\ell = \sum_{j,\, t_{\ell j}<20 \text{ min}} w(t_{\ell j}) S_j$$

where $t_{\ell j}$ is the driving time between the centroid of a block-group ℓ and preschool j and $\{w(t_{\ell j})\}$ are Gaussian decay weights defined as $w(t_{\ell j}) = \exp\left(\frac{-t_{\ell j}^4}{1000}\right)$. S_j constitutes an upper bound of the true number of seats per child as licensed capacity C_j reflects the size of the building rather than the actual number of available spots.[18] The resulting access measure A_ℓ is displayed on a map in Figure 2 panel (c). In spite of the higher density of children, preschoolers living in the center-rich area identified in panel (a) still have greater access to ECE services due to the concentration of centers in the area. In contrast, the less-educated, peripheral regions of Pittsburgh have fewer seats-per-child. Panel (d) confirms that this pattern identified for Pittsburgh holds for Pennsylvania at large. The range of the access metric shows that the vast majority of neighborhoods have fewer than one seat per preschooler, which implies that full enrollment cannot be met without some degree of supply expansion.

Figure D.1 in Appendix D shows that this heterogeneity in access is also reflected in children's attendance rate. While only around 50% of children go to preschools in census tracts with less than a quarter of college graduate women, this proportion climbs to almost 80% in tracts where the fraction of college graduates exceeds a half. The model covered in section 4 aims at discerning the extent to which these differences in access and enrollment can be attributed to parents' preferences or providers' costs.

These discrepancies in the geography of access to early education result from different patterns of preschools' dynamic decisions across neighborhoods. Figure 3 displays the 2018 state of centers that were rated STAR 1 in 2010. The left panel shows that in the poorest neighborhoods, almost 60% of these STAR 1 centers have exited the market during the panel, while this proportion falls below 30% for the richest census tracts. The right panel shows that conditional on being still in operation, less than 4% of centers serving low-income neighborhoods have upgraded to the highest rating, while almost 15% of centers in the richest census tracts have done so.

[18] The number of available spots is determined by the number of teachers, staff-to-children ratios and the number of classrooms. This shortcoming is common to most papers studying ECE access, and the only solution consists in surveying providers about their actual availability as in **?**.

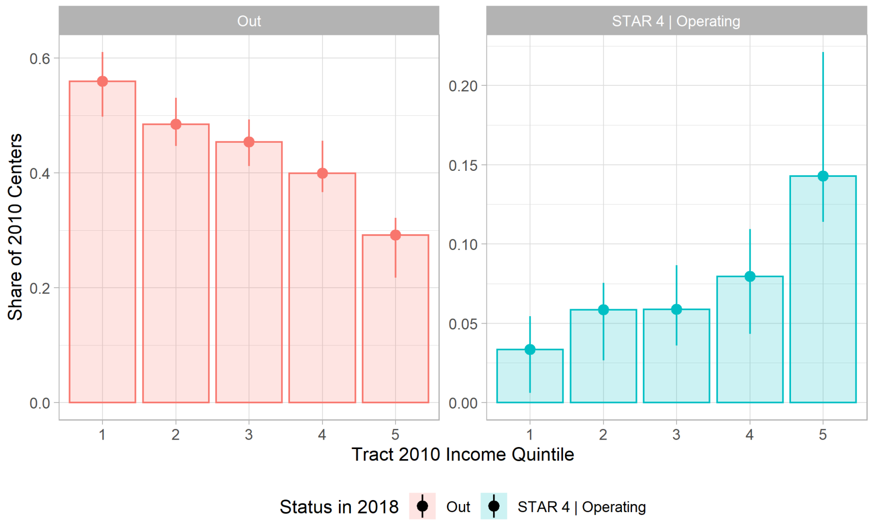


Figure 3: Exit and Upgrade Patterns Differ across Neighborhoods
Notes: Sample used in the Figure consists in centers rated STAR 1 in 2010. In 2018, these centers can either be out of the market or in one of the 4 ratings. Figure shows for each census tract income quintile, the proportion of these centers having exited on left panel, and conditional on still being in operation the proportion at rated STAR 4.

3.2 College-Educated Parents Value Preschool Quality

Table 1 shows that highly rated centers enroll more students and charge higher prices than lower quality preschools. To highlight the role of quality upgrades in those cross-sectional differences, I use the panel to compare the evolution of prices and enrollment between centers which have increased their quality between 2014 and 2018 and those who have not. Focusing on the sample of centers which were rated either STAR 1 or STAR 2 in 2014, I run the following regressions,

$$\Delta_0^t Y_{jc} = \beta sr_{jct} + \phi_c + \epsilon_j \tag{1}$$

where $\Delta_0^t Y_{jc}$ is the change in either daily price, full-time equivalent enrollment or private pay enrollment of preschool j operating in county c between school years 2014 and 2018, $sr_{jct} \in \{\text{STAR 1 or 2, STAR 3, STAR 4}\}$ stands for the STAR rating center j has achieved in 2018, and ϕ_c is a county fixed effect included in some of the specifications.

The results are presented in Table 2. The preferred specifications in columns 2, 4 and 6 include county fixed effects. Centers which upgrade to STAR 4 in the period considered increase their daily

Table 2: Price and Enrollment Changes of STAR 1 and STAR 2 Rated Preschools

	Δ Daily Price		Δ F.T.E Enrollment		Δ F.T.E Private Pay Enrollment	
STAR 3 in 2018	2.568***	4.989***	3.535***	14.255***	2.640***	9.640***
	(0.441)	(0.629)	(1.034)	(1.669)	(0.680)	(1.249)
STAR 4 in 2018	3.953***	7.138***	4.134***	15.511***	−0.422	6.043***
	(0.489)	(0.684)	(1.217)	(1.944)	(0.801)	(1.455)
County fixed effects		X		X		X
Observations	2,003	2,003	728	728	728	728

Note: *p<0.1; **p<0.05; ***p<0.01

rates on average \$7.1 more than centers remaining at their 2014 rating of STAR 1 or STAR 2. This is a large difference, corresponding to 19.8% of the average price of STAR 1 centers operating in 2018. Moreover, column 4 of Table 2 shows that this surge in daily prices for prekindergartens reaching the highest rating is accompanied by an increase in enrollment of 15.5 more children relative to centers remaining at lower quality. In addition, this enrollment increase is not limited to subsidized children, arguably less impacted by market prices; column 6 of Table 2 shows that private pay enrollment also rises more in centers reaching higher quality standards. The larger increase over this 5-year period of both prices and enrollment for upgrading preschools suggests that parents are willing to pay more for higher-rated preschools. To study whether prices increase more in highly educated neighborhoods following an increase to STAR 4, I estimate dynamic event studies with price as the outcome and increase to STAR 4 as the event of interest. I run separate regressions for centers serving in census tracts belonging to the first and fourth quartiles of share college educated women over the 2010-2018 period. To isolate the effect of an increase to STAR 4, I also control for the time since the upgrade to previous ratings. For each group of preschools, the estimated regression is

$$p_{jt} = \sum_{r \in \text{STAR } 2,3,4} \mathbf{1}_j(\text{upgrade } r). \sum_{\substack{\tau \geq -7 \\ \tau \neq -1}}^{8} \delta_\tau^{(r)} (t - t_j^{(r)} = \tau) + \phi_t + \psi_j + \epsilon_{jt}$$

Figure 4 shows the estimate coefficients $\{\delta_\tau^{(\text{STAR } 4)}\}$ for the schools serving the neighborhoods with the lowest and highest shares of college educated women. As preschools upgrade in different time periods, the coefficients $\delta^{(\text{STAR 4})}$ are estimated with cohort specific interactions following Sun and Abraham (2021). Coefficient on prices for centers in the least educated tracts are non significantly different from 0, even 5 years after the upgrade to STAR 4 occurred. In contrast, at that point in time centers located in highly-educated neighborhoods have increased their prices by almost \$3. This result suggests that college educated households are willing to pay for the higher prices charged by STAR 4 centers. The absence of significant price hikes from centers serving non college educated parents could reflect heterogeneity in preference, but also the policies in place to reward high-quality providers in these neighborhoods.

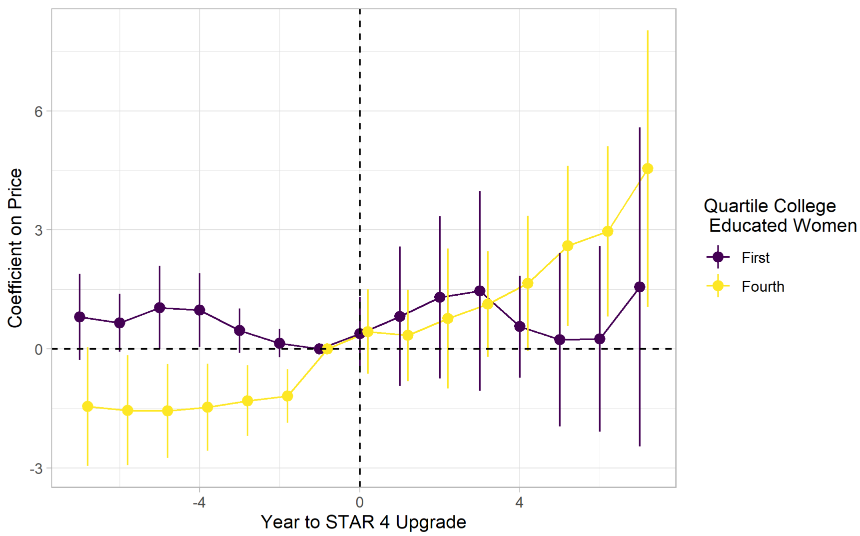


Figure 4: Price Trajectories around STAR 4 Upgrade

Notes: The effects of STAR 4 upgrades on prices are compared across centers in the lowest and highest quartile of census tracts in terms of share of college educated women. Coefficient estimates and 95% confidence intervals result from staggered treatment event-studies estimator.

3.3 Providers Respond to Financial Incentives to Operate at High-quality

As part of its QRIS program, Pennsylvania implements a set of policies aimed at encouraging providers to serve disadvantaged children at high-quality. The main supply-side component of these incentives is a schedule of add-on rates, which consists in a quality-specific bonus centers receive when serving a subsidized child. These bonuses are non-existent at the lowest quality tier for most of the period, and go up to almost 25% of the average daily price for the highest quality tier in 2018. For instance, a center rated STAR 4 in 2015 would receive from \$5 from the state for every subsidized child served *in addition* to the price covered by families' private payments and demand-side subsidies. This amount increased to \$7.5 in 2016 and \$9.2 in 2017. The full schedule of add-on rates between the years 2010 and 2018 can be seen in Figure B.1 in Appendix B. The general tendency over the sample period has been to increase financial incentives to serve at STAR 3 and STAR 4 ratings relative to STAR 2 and STAR 1.

Changes in the generosity of add-on rates vary the portion of revenue that providers receive from the state. In addition, centers serving many subsidized students benefit more from upgrading to a higher rating under a generous add-on schedule. I use this variation to test whether centers respond to financial incentives when deciding whether to change quality rating across years and estimate a conditional logit model. The sample includes all centers in all periods except the one where they exit. The alternatives are the 4 quality ratings. To capture costs of upgrading and the variation in the desirability of operating at each rating that is not due to the add-on schedule, the model also includes choice-specific year fixed effects and rating fixed effects. The variation isolated by the model is that between two centers deciding at which quality to operate next period and who expect different levels of additional revenue from the state either because the add-on schedule changed or because they tend to serve different populations. In this specification, the probability that a center j rated s chooses rating r at t is:

$$\mathbb{P}(r|C_{jst}) = \frac{\exp\left(\beta_r x_{jt} + \phi_{rt} + \mu_{rs}\right)}{\sum_{r'} \exp\left(\beta_{r'} x_{jt} + \phi_{r't} + \mu_{r's}\right)} \tag{2}$$

Table 3 presents estimates and average marginal effects from specifications without year fixed effects, without preschool characteristics, and with preschool price and capacity controls. Resulting estimates show a consistent, positive impact of additional revenue from the state on quality rating choices. The average marginal effect of additional \$100 of daily revenue (which would for instance result from a \$5 add-on rate applied to a center serving 20 subsidized children) on the probability of upgrading from STAR 3 to STAR 4 ranges between 2.1 and 4.1 percentage points depending on the specification. The average upgrade rate of STAR 3 schools over the sample is 12.6%, which implies that the effects are substantial. The fact that centers respond to financial incentives supports the profit-maximizing assumption on providers' behavior, and suggests that policies working through centers' revenues can impact the structure of the early education market.

Table 3: Upgrade Choice and Add-on Rates

	(1)	(2)	(3)
Additional Subsidy Revenue (USD 100)	0.354	0.276	0.183
	(0.045)	(0.04)	(0.043)
STAR Rating fixed effects	X	X	X
Year fixed effects		X	X
Preschool Controls			X
Absolute Marginal Effects			
Upgrade from 1 to 2	0.022	0.017	0.011
Upgrade from 2 to 3	0.035	0.027	0.018
Upgrade from 3 to 4	0.041	0.032	0.021
Observations	27023	27023	27023

Notes: Estimates from a multinomial logit model of next year STAR rating choice on expected revenue from subsidized students. The sample consists of all preschools in all years excluding the year in which they exit. Expected revenue from subsidized student is computed as the average subsidized enrollment of the center over the sample multiplied by the add-on rate paid by the state of Pennsylvania for each subsidized student served. Subsidized enrollement in each year is recovered from parents choices predicted by preschool choice model. Center controls included in column 3 are licensed capacity and posted price.

4 An Empirical Model of the Market for Preschools

This section presents the industry model describing evolution of the preschool market. Compared to previous models of education providers, this paper relies on a dynamic model of the supply side, as entry, exit and quality investments are key decisions through which centers can respond to a changing environment.[19] The main innovation compared to most of the literature on empirical dynamic games is to jointly model a static equilibrium of families' choices and providers' prices every period, and a dynamic equilibrium across periods.[20] The need for this complex modeling strategy arise from the counterfactual policies of interest in the early education market. For instance, subsidies to parents are central to policy discussions on ECE reform. A change in the schedule of these subsidies would affect the profits centers derive from serving certain neighborhoods in the short-run, and impact the distribution of centers across space in the long-run. A dynamic game framework with reduced-form flow profits would be poorly-suited to evaluate the consequence of such a policy.

The first part of this section presents the static model describing parents' preschool choices and centers' pricing decisions. Consistent with the evidence presented in the previous section, parents

[19]Neilson (2013) and Allende (2019) model quality responses as static. Singleton (2019) and Dinerstein et al. (2020) study entry and exit of schools but in a static game.

[20]The closest to this modeling approach is Sweeting (2013) in his study of radio stations. The main differences between our frameworks are that radio stations do not compete on prices, and are not spatially differentiated.

are endowed with heterogeneous preferences for providers characteristics and centers are assumed to behave like profit-maximizing firms. The second part of the section covers the model used for providers' dynamic decisions.

4.1 Static Demand and Supply of Preschool Education

Market environment. Following recent papers using structural models to study education markets, in particular (Neilson, 2013; Allende, 2019), families' choices of preschools are modeled using a spatially differentiated demand framework. I split Pennsylvania into geographically distinct areas. Each area m contains a fixed set of geographic nodes $\mathcal{L}_m$ in which families reside. I assume that parents from any node in $\mathcal{L}_m$ can send their child to any preschool within m.[21] A market is defined as a geographic unit (m) $\times$ school year (t). Each family i from node ℓ in school-year t is assumed to have one preschool-aged child and can have two observational types: $e \in \{H, L\}$ representing whether the mother has a college degree or not, and income y_{it} drawn from a location $\times$ type contingent distribution $y_{it} \sim F^y_{e\ell t}(.)$. Each node $\ell \in \mathcal{L}_m$ at time t is then characterized by a geographic location, a population count of 3 and 4 years old $N_{\ell t}$, a share of preschoolers' families with college-educated mothers, and two household income distributions $\left(F^y_{L\ell t}, F^y_{H\ell t}\right)$. This structure has the advantage of capturing the role of distance in a granular way without requiring micro-level data on families' addresses and early education arrangements.

Demand Subsidies. Section 2 outlined the set of incentives schemes implemented in Pennsylvania as part of the state's QRIS. Families with income below an eligibility threshold benefit from subsidies to help pay for early education. These demand side subsidies are controlled by two policy variables depicted in Figure B.2: a co-payment schedule $c_t(y)$, and a Maximum Child Care Allowance $MCCA_{mt}$. These policies determine the price subsidized parents with income y_{it} would actually be paying to send their child to a preschool with a market price of p_{jt}:

$$p_{ijt}(y_{it}) = \max\{p_{jt} - MCCA_{mt}, c_t(y_{it})\} \tag{3}$$

Product space and preferences. In each market mt operates a set of preschools $\mathcal{J}_{mt}$. A preschool j at time t is characterized by a location, a price charged for full-day education services

[21] I consider fairly large markets, as seen for instance on the Pittsburgh map in Figure 2. This reduces the number of artificial boundaries assumed impassable for parents. These large markets are constructed from aggregating school districts together, and for large urban areas distinctive geographic features are used to separate the city into distinct markets, such as rivers in the case of Pittsburgh.

p_{jt}, and a STAR quality rating sr_{jt}. In addition, preschool j in year t is characterized by a fixed unobserved quality component ξ_j, a time specific demand shock ξ_t, and a transitory demand shock $\Delta\xi_{jt}$. These terms capture characteristics of a center that are not observed by the econometrician but are known by parents when choosing a preschool. The fixed component ξ_j could represent fixed features of the location preschool j operates in. I allow families to have heterogeneous across education types and income levels. The utility derived by household i of education e living in neighborhood ℓ from sending her child to preschool j located at a $d_{j\ell}$ drive time from ℓ is given by:

$$\begin{aligned} u_{i\ell ejt} = & \underbrace{\beta^{(sr)}\mathbf{1}\{sr_{jt} = sr\} + \xi_t + \xi_j + \Delta\xi_{jt}}_{\delta_{jt}} \\ & + \underbrace{\beta_H^{(0)}\mathbf{1}\{e_i = H\} + \beta_H^{(sr)}\mathbf{1}\{e_i = H\}\mathbf{1}\{sr_{jt} = 3,4\} + \left(\alpha_1^e + \frac{\alpha_2^e}{y_{it}}\right) p_{ijt}(y_{it}) + \lambda^e d_{\ell j}}_{\mu_{e,\ell,j}(y_{it})} + \varepsilon_{ijt} \end{aligned} \tag{4}$$

Income y_{it} is allowed to influence household i price sensitivity as in Berry et al. (1995, 1999) and ε_{ijt} is a family specific taste shock for preschool j at t[22]. Assuming these shocks are Extreme Value Type I distributed, and denoting by $\{sub, nosub\}$ the subsidy status of families, the share of education e households living in neighborhood ℓ opting for preschool j is given by:

$$\begin{aligned} s_{\ell jt}^e = & \int_0^\infty \frac{\exp(\delta_{jt} + \mu_{e,\ell,j}(y_{it}))}{1 + \sum_{k\in\mathcal{J}_{mt}} \exp(\delta_{kt} + \mu_{e,\ell,k}(y_{it}))} dF_{e\ell t}^y \\ \approx & \sum_n w_{\ell n}^{e,sub} \underbrace{\frac{\exp(\delta_{jt} + \mu_{e,\ell,j}(y_n))}{1 + \sum_{k\in\mathcal{J}_{mt}} \exp(\delta_{kt} + \mu_{e,\ell,k}(y_n))}}_{s_{\ell njt}^{e,sub}} + \sum_n w_{\ell n}^{e,nosub} \underbrace{\frac{\exp(\delta_{jt} + \mu_{e,\ell,j}(y_n))}{1 + \sum_{k\in\mathcal{J}_{mt}} \exp(\delta_{kt} + \mu_{e,\ell,k}(y_n))}}_{s_{\ell njt}^{e,nosub}} \end{aligned} \tag{5}$$

where the last row corresponds to a numerical integration step.[23] The aggregate market share of preschool j at time t then results from summing over neighborhoods ℓ, education e, and income n nodes:

[22] This functional $\frac{1}{y_i}$ can be seen as first-order Taylor approximation of a log utility.

[23] Gauss-Legendre quadrature is used to approximate the integrals above, see Appendix E1 for additional details.

$$s_{jt}(\boldsymbol{p}_t, \boldsymbol{sr}_t, \boldsymbol{\xi}_t) = \sum_{\substack{\ell \in \mathcal{L}_m \\ e=L,H}} w^e_{\ell t} \sum_{\substack{\tau \in \{sub, nosub\} \\ n \in y(\tau)}} w^{e,\tau}_{\ell n t} s^{e,\tau}_{\ell n j t}(\boldsymbol{p}_t, \boldsymbol{sr}_t, \boldsymbol{\xi}_t) \tag{6}$$

Several remarks can be made on this framework of preschool choice. I do not model all the potential substitutes for prekindergarten. Parental and relative care or family-based day care services are bundled in households' outside option. However, accounting for heterogeneity in household preferences captures the fact that families of different socioeconomic backgrounds might have different alternative arrangements available. This parsimonious approach also captures the main dimensions of competition between preschools. Centers can be differentiated along observed and unobserved quality components, and do not face the same local demand depending on their location. Potential capacity constraints faced by households are assumed away. Figure C.2 in Appendix C shows that these constraints are unlikely to be binding for various conservative measures of centers' effective capacities. A limitation of this modeling strategy is that the preference parameters do not take into account parents' labor market decisions. This model captures the various trade-offs faced by families when choosing among early education arrangements, but a *ceteris paribus* increase in the opportunity cost of parental time could for instance translate in an increase in $\beta^{(0)}_H$. In this paper, I take the labor market as exogenous to instead focus on the structure of the early education market.

Marginal Costs of Early Childhood Education. I take advantage of the explicit modeling of the demand system detailed in the previous paragraph to back out marginal costs from the observed pricing behavior of early education providers. In the static part of the model, quality rating is taken as given and centers compete on prices. I assume that preschools are profit maximizing firms in Bertrand competition.[24] The government provides preschools operating at quality sr with financial support $\kappa(sr)$ for subsidized child served. For a given quality rating sr_{jt} and unobserved shocks to demand and cost, preschool j chooses price p_{jt} to maximize its variable flow profits given by

$$\pi(p_{jt}, \{s^\tau_{jt}\}_\tau, sr_{jt}) = \sum_{\tau \in \{sub, nosub\}} \pi^\tau_{jt} = \sum_{\tau \in \{sub, nosub\}} N_{mt} w^\tau_{mt} s^\tau_{jt} \left(p_{jt} + \kappa^\tau(sr_{jt}) - mc_{jt}\right) \tag{7}$$

[24]Although some centers are non-profits, the majority of preschools in the data are for profits. In addition, without further assumptions I cannot distinguish preschools with low marginal costs from preschools with an altruistic objective which places weight on enrollment. I therefore treat all centers as pure profit maximizers.

Under these assumptions, optimal price for preschool j is equal to its marginal cost mc_{jt} plus a markup η_{jt} defined as follows

$$mc_{jt} = p_{jt} - \underbrace{\left(- \left[\frac{\partial s_{jt}}{\partial p_{jt}} \right]^{-1} \left(s_{jt} + \sum_{\tau \in \{sub, nosub\}} w_{mt}^{\tau} \frac{\partial s_{jt}^{\tau}}{\partial p_{jt}} \kappa^{\tau}(sr_{jt}) \right) \right)}_{\text{markup } \eta_{jt}(\boldsymbol{p}_t, \boldsymbol{s}_t; \kappa)} \tag{8}$$

The expression above highlights the impact of the supply side policy κ on centers' pricing decisions. A higher level of financial support from the state leads high-quality centers serving more subsidized students to lower their markup above marginal cost, because public funds act as substitutes to higher prices.

4.2 Entry, Exit and Quality Upgrades

State space, actions and timing. The previous paragraphs presented the model representing families' choices over preschools every year and providers pricing decisions. In addition, crucial dimensions of the ECE market are the availability and quality of centers nearby, determined by the entry, exit, and upgrade choices of providers over the years. These decisions are based on families' demographics, providers characteristics including quality ratings and fixed and transitory demand and cost shocks, and the spatial structure of drive times between families and providers. The state $\mathcal{M}_{xjt}$ represents all of this public payoff-relevant information at time t perceived by preschool j operating in own-state $x \in \{\text{Out}, \text{STAR 1}, \text{STAR 2}, \text{STAR 3}, \text{STAR 4}\}$. Every period, actions available to an incumbent provider are $\mathcal{A}(\mathcal{M}_{xjt}) = \{\text{Out}, \text{STAR 1}, \text{STAR 2}, \text{STAR 3}, \text{STAR 4}\}$. Potential entrants are assumed to be tied to a specific location in which they can choose to operate, and their actions boil down to choosing between entering at STAR 1 or staying-out.[25] Figure F.4 in Appendix F presents the frequencies of each of these actions by state in the data. I assume that information realization, decisions, and payoffs follow the timing below:

1. $\mathcal{M}_{xjt}$ is realized
2. Fixed costs of operating at the current state x, $C(x_{jt})\psi_C$, are paid
3. Private action-specific payoff shocks are realized $\epsilon(a)$. Own-state for next period is chosen, potentially resulting in a scrap value, transition or entry cost $W(a_{jt}, \mathcal{M}_{xjt})\boldsymbol{\psi}_W$

[25] This modeling choice differs from other spatial competition games such as Seim (2006) or more recently Caoui et al. (2022). For child care centers, suggestive evidence supports that providers are more tied to a specific neighborhood than the video rentals and supermarket industries studied by these papers.

4. Variable profits $\breve{\pi}(\mathcal{M}_{xjt})$ from static price-competition are realized

5. Next period state is determined by j's chosen action a_{jt}, j competitors' actions and aggregate state transitions

The flow payoffs π from steps 3 of t to step 2 of period $t+1$ $\Pi(a_{jt}\mathcal{M}_{xjt}) + \boldsymbol{\psi}_\epsilon \epsilon(a_{jt})$ are defined as:

$$\Pi(a_{jt}\mathcal{M}_{xjt}) + \boldsymbol{\psi}_\epsilon \epsilon(a_{jt}) = \pi(\mathcal{M}_{xjt}) - W(a_{jt}, \mathcal{M}_{xjt})\boldsymbol{\psi}_W - \beta C(a_{jt})\psi_C + \boldsymbol{\psi}_\epsilon \epsilon(a_{jt}) \tag{9}$$

Equilibrium Concept. I focus on stationary Markov-Perfect Bayesian Nash Equilibria (MPE). I denote the strategy of a firm in state $\mathcal{M}_{xjt}$ with payoff shocks ϵ_{jt} as $\mathcal{P}_j(\mathcal{M}_{xjt}, \epsilon_{jt})$, and the corresponding strategy profile as $\mathcal{P} = \{\mathcal{P}_j(\mathcal{M}_{xjt}, \epsilon_{jt})\}_{j\in\mathcal{J}}$.[26] The ex-ante value function of preschool j in state $\mathcal{M}_{xjt}$ given a strategy profile $\mathcal{P}$ can be written as:

$$V^{\mathcal{P}}(\mathcal{M}_{xjt}) = \mathbb{E}_{\epsilon_t} \max_{a_{jt}} \{v^{\mathcal{P}}(a, \mathcal{M}_{xjt}) + \boldsymbol{\psi}_\epsilon \epsilon(a_{jt})\} \tag{10}$$

where $v^{\mathcal{P}}(a, \mathcal{M}_{xjt})$ is choice specific value function corresponding to action a in state $\mathcal{M}_{xjt}$, defined as:

$$v^{\mathcal{P}}(a, \mathcal{M}_{xjt}) = \Pi(a, \mathcal{M}_{xjt}) + \beta \int V^{\mathcal{P}}(\boldsymbol{\mathcal{M}_{\tilde{x}jt+1}}) g(\boldsymbol{\mathcal{M}_{\tilde{x}jt+1}} | a, \mathcal{P}_{-j}, \mathcal{M}_{xjt}) d\boldsymbol{\mathcal{M}_{\tilde{x}jt+1}} \tag{11}$$

The second term corresponds to the future value of taking action a in state $\mathcal{M}_{xjt}$ given a strategy profile $\mathcal{P}$ and is denoted by $FV(a, \mathcal{M}_{xjt}, \mathcal{P})$. Making the dependence on dynamic parameters explicit, the choice-specific value function can be written as

$$v^{\mathcal{P}}(a, \mathcal{M}_{xjt}; \boldsymbol{\psi}) = \Pi(a, \mathcal{M}_{xjt}; \boldsymbol{\psi}) + FV(a, \mathcal{M}_{xjt}, \mathcal{P})$$

The transition kernel $g(\boldsymbol{\mathcal{M}_{\tilde{x}jt+1}} | a_{jt}, \mathcal{P}_{-j}, \mathcal{M}_{xjt})$ incorporates transitions of exogenous variables relevant for profits such as demographics and i.i.d. transitory demand and costs shocks, and also

[26] The focus on stationary equilibria requires that players' strategies are not indexed by t. While the environment centers compete in evolves with time, the relevant information on time-varying factors is already captured in $\mathcal{M}_{xjt}$ and players' expectations. The exception are the changes in the policy environment described in appendix B, for which expectations are not modeled. Instead, I assume that centers are myopic and believe that next periods' policy environment will prevail forever.

expectations over other players actions. A MPE is a strategy profile $\mathcal{P}$ satisfying

$$\mathcal{P}^*(\mathcal{M}_{xjt}, \epsilon_{jt}) = \arg\max_{a_{jt}} \{v^{\mathcal{P}}(a, \mathcal{M}_{xjt}) + \boldsymbol{\psi}_\epsilon \epsilon(a_{jt})\}$$

Players' conditional choice probabilities (CCP) in equilibrium $\mathbb{P}$ are obtained by integrating $\mathcal{P}^*$ over private payoff shocks. Assuming Extreme Value Type I distribution for the action specific payoff shocks ϵ, the CCP corresponding to provider j's best response to the strategies $\mathcal{P}$ can expressed as

$$\mathbb{P}^{\mathcal{P}}(a_{jt}|\mathcal{M}_{xjt}; \boldsymbol{\psi}) = \frac{\exp\left(\frac{v^{\mathcal{P}}(a, \mathcal{M}_{xjt}; \boldsymbol{\psi})}{\boldsymbol{\psi}_\epsilon}\right)}{\sum_{a'} \exp\left(\frac{v^{\mathcal{P}}(a', \mathcal{M}_{xjt}; \boldsymbol{\psi})}{\boldsymbol{\psi}_\epsilon}\right)} = \frac{\exp\left(\frac{\Pi(a, \mathcal{M}_{xjt}; \boldsymbol{\psi}) + FV(a, \mathcal{M}_{xjt}, \mathcal{P})}{\boldsymbol{\psi}_\epsilon}\right)}{\sum_{a'} \exp\left(\frac{\Pi(a', \mathcal{M}_{xjt}; \boldsymbol{\psi}) + FV(a', \mathcal{M}_{xjt}, \mathcal{P})}{\boldsymbol{\psi}_\epsilon}\right)} \tag{12}$$

Curse of dimensonality. The model outlined above is difficult to solve in practice due to the dimensionality of the state space. Public information available to providers includes multiple continuous variables, such as demographics and providers' unobservable characteristics, making the state space high-dimensional. In addition, several urban markets feature many dozens of players, further complicating the game. This complexity in turn makes the value function $V^{\mathcal{P}}$ and transition kernel g intractable. To solve this dynamic game, I use a recursive procedure that iterates on centers' strategies, discretizes the state space and approximates the value function to make the computation tractable. Section 5.2 details the solution method and estimation approach used in this paper.

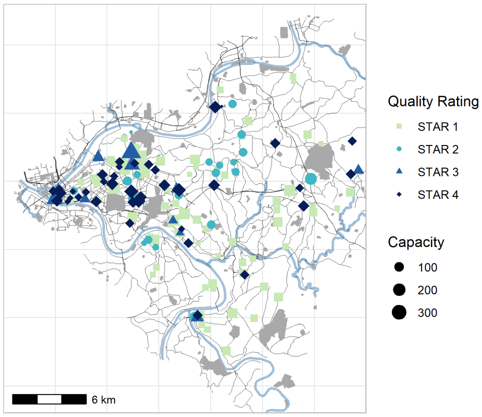
5 Estimation and Solution Method

Estimation of the model proceeds in two main steps. The first stage consists in estimating the static parameters of the preschool choice and marginal costs models. The second stage estimates fixed costs, entry and quality investments costs governing the dynamic decisions of providers. The next section outlines the key aspects of these two estimation stages, while details can be found in Appendix E and F.

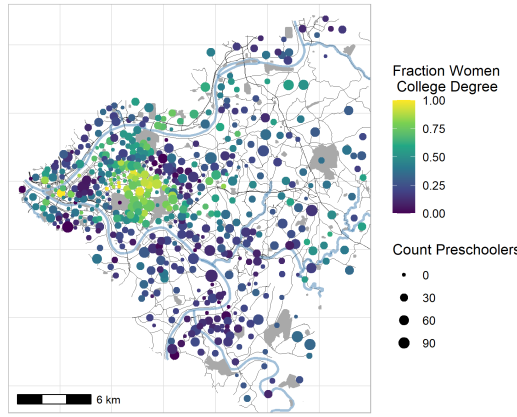
5.1 Estimation and Identification of the Static Model

Spatial Structure of Markets. Figure 5 shows the geographic structure use in estimation for the market labeled "East Pittsburgh" in 2018. Panel (a) displays the distribution of centers families in this market can choose from, and their quality rating. Panel (b) shows the distribution of the nodes $\mathcal{L}_m$ in the market, their population of preschoolers and their share of college educated women.

Nodes are taken as centroids of census block groups. To reduce border effects, the market includes a buffer of nodes beyond the border in which preschools are contained.[27]



(a) Centers by Quality Rating



(b) Preschoolers by Education

Figure 5: Product and Consumer Static Model

Notes: Left panel: Centers operating in 2018 in the Pittsburgh-East geographic market. Right panel: Each dot represents the geographic centroid of a census block group. College Degree comprises 4 year college degrees, professional, advanced and graduate degrees. Preschoolers stands for children aged 3 and 4. Demographic variables come from the ACS 2014-2018 block-group level 5-year estimates.

Addressing Missing Market Shares In Demand Estimation. For clarity, the exposition of the demand model does not explicit the fact that a subset of centers do not report their enrollment. In estimation, this is accounted for by grouping this subset of non-reporting centers into one common group of products, and by matching an estimate of their combined market share.[28] Their location, and prices still create differentiation between these centers, but only one unobserved quality term is recovered for the entire group of products. This method, inspired by Elliott et al. (2021), allows me to preserve the information contained in the spatial distribution and prices of centers in spite of the absence of some of the product shares. Appendix E5 provides details on the approach.

Static Preference and Marginal Cost Parameters. The parameters of the demand model consist in common (linear) preferences for school ratings $\boldsymbol{\theta}_1 = \{\beta^{(sr)}\}_{sr\in STAR\,1,\ldots,4}$ and in family type specific (non-linear) parameters $\boldsymbol{\theta}_2 = (\beta^{(0)}_H, \beta^{(sr)}_H, \{\alpha^e_1, \alpha^e_2, \lambda^e\}_{e\in L,H})$. Lastly, the parameters of the demand system also include the fixed unobserved quality $\boldsymbol{\xi}$ and transitory demand shocks $\boldsymbol{\Delta\xi}$.

[27]A similar strategy is used by Allende (2019).

[28]This estimate is obtained by aggregating the predictions of a machine learning model taking as predictors demographics, centers characteristics, and local preschool attendance rates as reported in the ACS.

The marginal cost is assumed to depend linearly on observed preschool characteristics $\boldsymbol{X}_{jt}$ and unobserved residual cost components. The observed preschool characteristics include in particular the star ratings of the center at time t, but also observables of the neighborhood such as demographics and density. Similar to the residuals in the common component of utility, unobserved cost shocks are decomposed in a time fixed effect ω_t, a preschool fixed effect ω_j, and a transitory component $\Delta\omega_{jt}$ as follows:

$$mc_{jt} = \omega_t + \omega_j + X_{jt}\boldsymbol{\theta_3}' + \Delta\omega_{jt} \tag{13}$$

Estimator and Identification. The static parameters $(\boldsymbol{\theta}, \boldsymbol{\xi}, \boldsymbol{\omega}, \boldsymbol{\Delta\xi}, \boldsymbol{\Delta\omega})$ are estimated using generalized method of moments (GMM) following the standard approach for differentiated products demand covered in Berry et al. (1995) and Conlon and Gortmaker (2020). The principal concern at this stage of the estimation is the endogeneity of prices and market shares. As the model presented in the previous section includes center fixed effects $(\boldsymbol{\xi}, \boldsymbol{\omega})$, estimation requires instruments orthogonal to the transitory shocks $\boldsymbol{\Delta\xi}, \boldsymbol{\Delta\omega}$. Details of the instruments are presented in appendix E, but I outline the main assumptions underlying the exclusion restriction in this paragraph. The instruments are constructed following the approach proposed by Gandhi and Houde (2019). It consists in creating measures of how isolated products are in the characteristics space. Market structure is endogenous in the model, which implies that key product characteristics such as the number and quality ratings of competitors are also endogenous. However, the *timing* of the dynamic game by construction makes the transitory shocks orthogonal to the market structure in period t. As a consequence, for the sake of the static estimation, competitors characteristics except for prices can be used as instruments. Denoting demand and supply side instrumental variables as Z_D, Z_S, I construct the following moments used in a GMM estimator

$$g^{IV}(\boldsymbol{\theta}) = \frac{1}{N}\begin{bmatrix} Z_D' & 0 \\ 0 & Z_S' \end{bmatrix}\begin{bmatrix} \Delta\xi(\boldsymbol{\theta}) \\ \Delta\omega(\boldsymbol{\theta}) \end{bmatrix}$$

Additional Moments. In addition to the instrumental variable moments, distinctive features of the data provide additional moments that can be used in estimation. First, centers report not only their aggregate market shares, but also how their enrollment is distributed between subsidized and non-subsidized children. As subsidy status is mostly based on income, and income distributions differ by geography and educational attainment, this is likely to provide variation useful for estimating the type-specific preference parameters. The following moments are constructed for each market m, t and added to the GMM objective:

$$g_{mt}^{nosub}(\boldsymbol{\theta}) = \frac{1}{\mathcal{J}_{mt}} \sum_{j \in \mathcal{J}_{mt}} (s_{jt}^{nosub}(\boldsymbol{\theta}) - s_{jt}^{nosub,data})$$

I also use microdata on families'choices from the NSECE to form additional moments linking demographics to characteristics of early education arrangements. Additional details about the static estimation can be found in appendix E.

5.2 Solution Method for the Dynamic Game

Estimating the dynamic parameters requires bypassing the curse of dimensionality. Two main approaches have been used in the literature on empirical dynamic games. One method builds on the workhorse technique for dynamic discrete choice models and relies on exploiting *finite dependence* to eliminate the need to compute value functions in estimation.[29] While the presence of exit as a terminal action makes the framework described above a possible candidate for this approach, in practice this method hinges on precise estimates of exit probabilities. I rely instead on a second approach implemented by Sweeting (2013) which combines a parametric approximation of the value function with a variant of *nested pseudo likelihood* (NPL) estimator of Aguirregabiria and Mira (2007, 2010). This estimator only relies on CCP estimates to provide a starting point, and requires the final strategy profile of players to be a MPNE. Imposing these additional equilibrium restrictions also makes this estimator more efficient than two-stage approaches.[30]

Parametric Approximation of the Value Function using neural networks. The Bellman equation (10) can be expressed in terms of the expected flow profits under $\mathbb{P}$, denoted by $\bar{\Pi}(\mathcal{M}_{xjt}; \boldsymbol{\psi}, \mathbb{P})$

$$V^{\mathbb{P}}(\mathcal{M}_{xjt}) = \bar{\Pi}(\mathcal{M}_{xjt}; \boldsymbol{\psi}, \mathbb{P}) + \beta \mathbb{E}_{\mathbb{P}} \left[V^{\mathbb{P}}(\boldsymbol{\mathcal{M}_{\tilde{x}jt+1}}) \right] \tag{14}$$

where $\bar{\Pi}(\mathcal{M}_{xjt}; \boldsymbol{\psi}, \mathbb{P}) = \sum_a \mathbb{P}(a, \mathcal{M}_{xjt}) \left(\pi(\mathcal{M}_{xjt}) - W(a, \mathcal{M}_{xjt})\boldsymbol{\psi}_W - \beta C(a)\psi_C + \psi_\epsilon[\gamma - \log \mathbb{P}(a, \mathcal{M}_{xjt})]\right)$

[29] This is the approach used by Caoui et al. (2022).

[30] Bugni and Bunting (2021) propose an alternative estimator (k-MD) with desirable properties but this estimator can imply a heavy computational burden in games with a large number of players such as in this paper. A possible alternative would be the estimator developed by Dearing and Blevins (2019).

Solving for firms' equilibrium strategies requires evaluating their ex-ante value functions $V^{\mathbb{P}}$. The state space is high-dimensional, which makes the exact computation of $V^{\mathbb{P}}$ intractable. I follow Sweeting (2013) and assume that the ex-ante value function $V^{\mathbb{P}}$ can be approximated by a linear combination of functions of the state. The main challenge to this approach is that forming this approximation requires solving the equilibrium of centers pricing game a very large number of times. I use a neural network trained to directly compute centers' variable profits in each state to make this computation feasible. The static equilibria are only computed once on a large number of states to train the network, and during estimation the network is used to get predicted variable profits and form the basis of approximation for the value function. I detail the approach in the paragraph below.

I assume that $V^{\mathbb{P}}$ can be computed as a linear combination of K basis functions $\mathbf{\Phi}$. Intuitively, this basis should be a low-dimensional representation of the state space containing functions that are particularly relevant for flow profits Π, such as variable profits π. In principle, $\mathbf{\Phi}$ should be evaluated on every possible state, but this is not feasible due to the dimension of the state space. I instead discretize the state space in a grid of N points. This grid includes all the states ever visited in the data. To address the concern of out-of-sample validity of the value function approximation, I augment the data states with perturbed duplicates of these states where market structure, demographics, and transitory shocks are randomly changed. For each states $\mathcal{M}_{xjt}$ in this grid, I simulate a set of H potential next states $\mathcal{H}(\mathcal{M}_{xjt})$ which are used to compute the expectation of the value function next period.[31] Computing this expectation requires evaluating $\mathbf{\Phi}$ on each of the states in in $\{\mathcal{H}(\mathcal{M}_{xjt}), \mathcal{M}_{xjt} \in N\}$. I use a neural network to speed up this computation. The network is trained on the grid of N states to approximate the Nash Bertrand pricing equilibrium function $\mathcal{M}_{xjt} \to \pi(\mathcal{M}_{xjt})$.[32] It provides a fast and reliable solution to compute variable profits in each state of $\mathcal{H}(\mathcal{M}_{xjt})$.[33] I then use the output of the trained network $\tilde{\pi}(\boldsymbol{\mathcal{M}_{\tilde{x}jt+1}})$, and its hidden layers as elements of $\mathbf{\Phi}$. A detailed description of the construction of $\mathbf{\Phi}$ can be found in appendix F. With this approach, the value function and Bellman equation can be expressed only in terms of

[31] In estimation and counterfactuals I take $H = 500$.

[32] On its own, training the network already requires calculating the solution of the Nash Bertrand pricing equilibrium on the N states of the grid. Appendix E8 details how these computations are made efficient.

[33] I found neural network performed better at this task than a LASSO on a rich set of basis, see Appendix F4.

the vectors $\boldsymbol{\rho} \in \mathbb{R}^K, \bar{\Pi}(\mathbb{P}) \in \mathbb{R}^N$ and real matrices $\boldsymbol{\Phi} \in \mathbb{R}^{N \times K}, \mathbb{E}_{\mathbb{P}}[\boldsymbol{\Phi}] \in \mathbb{R}^{N \times K}$ as

$$V^{\mathbb{P}}(\mathcal{M}_n) \approx \sum_k \rho_k \phi_k(\mathcal{M}_n) = \boldsymbol{\Phi}\boldsymbol{\rho}$$

$$\boldsymbol{\Phi}\boldsymbol{\rho} = \bar{\boldsymbol{\Pi}}(\boldsymbol{\psi}, \mathbb{P}) + \beta \mathbb{E}_{\mathbb{P}}[\boldsymbol{\Phi}]\boldsymbol{\rho}$$

As $N >> K$, the Bellman equation above can be used to retrieve $\boldsymbol{\rho}$ for a given CCP profile $\mathbb{P}$ using ordinary least squares:[34]

$$\hat{\boldsymbol{\rho}}(\boldsymbol{\psi}, \mathbb{P}) = (\boldsymbol{\Phi} - \beta \mathbb{E}_{\mathbb{P}}[\boldsymbol{\Phi}])'(\Phi - \beta \mathbb{E}_{\mathbb{P}}[\boldsymbol{\Phi}])]^{-1}(\boldsymbol{\Phi} - \beta \mathbb{E}_{\mathbb{P}}[\boldsymbol{\Phi}])'\bar{\boldsymbol{\Pi}}(\boldsymbol{\psi}, \mathbb{P}) \tag{15}$$

Iterative Solution Method. The parametric policy iteration (PPI) builds on the value function approximation to solve the dynamic game by iterating over the CCP profile. In estimation, the dynamic parameters $\boldsymbol{\psi}$ are unknown and are estimated at the same time as the game being solved. The procedure is the following

1. Estimate first-stage CCP $\hat{\mathbb{P}}$ and compute $\boldsymbol{\Phi}$, $\mathbb{E}_{\hat{\mathbb{P}}}[\boldsymbol{\Phi}]$ on all grid points N
2. Start from an initial guess of parameters $\boldsymbol{\psi}^0$ and CCP profile $\mathbb{P}^0 = \hat{\mathbb{P}}$

Given these initial objects, for each iteration $i : (\boldsymbol{\psi}^i, \mathbb{P}^i)$, the following steps are repeated until CCP profile parameters converge:

3. Compute the approximation weights $\hat{\boldsymbol{\rho}}(\boldsymbol{\psi}^i, \mathbb{P}^i)$ from (15)
4. Compute choice-specific future values $FV(a, \mathbb{P}^i) = \beta \mathbb{E}_{\hat{\mathbb{P}}}[\boldsymbol{\Phi}]\hat{\boldsymbol{\rho}}(\boldsymbol{\psi}^i, \mathbb{P}^i)$
5. Estimate new parameters $\boldsymbol{\psi}^{i+1}$ by maximum likelihood where probabilities are given by (12)
6. Compute the new CCP profile $\mathbb{P}^{i+1}$ given the new parameters $\boldsymbol{\psi}^{i+1}$ using (12)

Figure F.3 in Appendix proposed a graphical overview of the objects involved in the value function approximation, and of how this approach helps bypassing the curse of dimensionality. Additional details can be found in Appendix F. To perform counterfactuals, a similar iterative algorithm is used with the only difference being that $\boldsymbol{\psi}$ is known, so step 5. becomes unnecessary.

[34] In practice, I find that adding a small Ridge penalty stabilizes the solution method without significantly impacting the fit of the empirical Bellman equation.

Computing Expectations. Solving this dynamic game requires repeatedly integrating the value function over future states. While the linear approximation makes it feasible to evaluate the value function in a given state $\mathcal{M}_{xjt}$, the dimension of the state space and the number of players require an efficient method to compute expectations. I discretize the next period state space by drawing a set $\mathcal{H}(\mathcal{M}_{xjt})$ of H possible future states for each of the N current states. The neural network approximation to variable profits provides an efficient way to compute profits without solving for the static Bertrand pricing equilibrium in each of the $N \times H$ future states.[35] Given this grid, expectations are computed using weighted importance sampling:

$$\mathbb{E}_{(\mathbb{P}^i_j, \mathbb{P}^i_{-j})}[\phi_k(\mathcal{M}_n)] \approx \sum_{\mathcal{M}_h \in \mathcal{H}(\mathcal{M}_n)} \phi_k(\mathcal{M}_h) \frac{w(\mathcal{M}_h)}{\sum_{\mathcal{M}_{h'}} w(\mathcal{M}_{h'})} \text{ where } w(\mathcal{M}_h) = \frac{\mathbb{P}^i_j(\mathcal{M}_h)\Pi_{k \neq j}\mathbb{P}^i_k(\mathcal{M}_h)}{\hat{\mathbb{P}}_j(\mathcal{M}_h)\Pi_{k \neq j}\hat{\mathbb{P}}_k(\mathcal{M}_h)}$$

Multiple Equilibria in Estimation. Dynamic games such as the model described in the previous section can have multiple equilibria which may cause the NPL procedure to fail to converge (Pesendorfer and Schmidt-Dengler, 2010). To overcome this issue in estimation, I follow Sweeting (2013) and apply a variant of the NPL procedure proposed by Aguirregabiria and Mira (2010) where beliefs about competitors' actions remain fixed at their initial values given by the first-stage CCPs, i.e $\mathbb{P}^i_{-j} = \mathbb{P}^0_{-j} = \hat{\mathbb{P}}_{-j} \; \forall \; i, j$. This assumption on preschools' beliefs changes the algorithm to an iterative procedure on parameters and strategies only, and guarantees convergence.[36]

6 Empirical Results

This section discusses the results from the estimation of the static and dynamic parameters.

6.1 Static Parameters

Table 4 present the estimation results for the parameters governing the families choices of providers.[37] These estimates highlight two main differences between families with college and non-college educated mothers. First, college educated mothers value early education, and quality in particular, more than their non college educated counterpart. The linear parameters β^{sr} measuring the prefer-

[35] A similar approach is used by Gowrisankaran et al. (2022) who extrapolate profits of electricity producers based on a flexible regression.

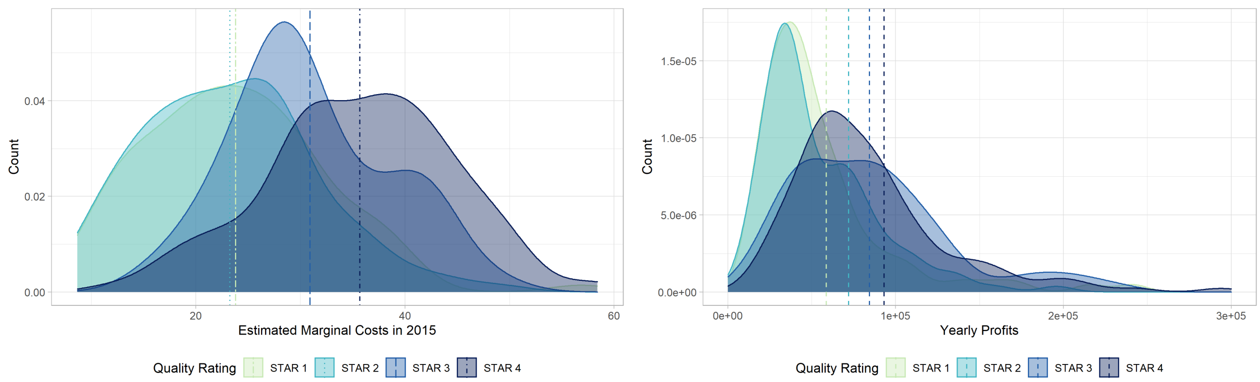
[36] In future versions of this paper, I could still use the full NPL procedure to check whether the procedure converges, and if so whether the estimated parameters would substantially change.

[37] The income heterogeneity coefficient on price for non college-educated parents hits an upper bound during estimation. The difficulty to estimate this coefficient likely comes from the fast that many of these households are partially shielded from price variation across products due to demand subsidies.

ence for higher ratings of non-college educated parents are close to 0. In contrast the high-quality premium $\beta_H^{(q)}$ for college-educated parents translates in marginal willingness to pay for STAR 3 and 4 ratings centered around \$2 and \$4 respectively. The difference between the willingness to pay for quality across demographics is consistent with the event studies in section 3 showing a larger price increase following a STAR 4 upgrade in neighborhoods with more college graduates. Second, non college-educated parents are less willing to travel then their college-educated counterparts. The marginal willingness to pay to avoid additional travel time of the first group are centered around \$4.6 while that of the second group are centered around \$1.2. This difference could capture either a wider breadth of nearby alternative arrangements available to non college-educated households, or a tendency for the college-educated households to rely on providers located near their workplace, further away from home.

Table 5 presents estimates for the key coefficients of providers marginal costs. The estimates show that a ceteris paribus increase in quality from STAR 1 to STAR 4, is associated with an increase in daily cost per child of almost \$6, which correspond to more than a fifth of the average daily price of centers. Estimates also suggest that centers serving a higher fraction of children aged 0 to 2 (infants and toddlers) have higher marginal costs, and that costs decrease with experience. This is in line with previous studies of early education supply highlighting the higher costs of serving younger children. Overall, the coefficients have the expected signs and the magnitudes of marginal costs are reasonable. A conversion in dollar per child of the cost estimates will be given when discussing the dynamic estimates, so that fixed costs can be incorporated in the calculation.

Figure 6 displays the recovered variable yearly profits for centers. In the current policy environment, it is on average both more costly per child and profitable overall to operate at the highest quality tiers, but this masks considerable heterogeneity. In addition, as the markups shown in appendix E suggest, the profits for the higher quality tiers originate in large parts from the state financial supports schemes, which allow these centers to charge lower markups.



(a) Marginal Costs (b) Variable Profits π

Figure 6: Static Estimation Results, Supply Side, 2015

Notes: Distribution of estimated marginal costs and resulting yearly variable profits from the static supply model. Profits do not include fixed costs yet, but incorporate financial support received from the state in the form of add-on rates.

Table 4: Demand Estimates

Coefficients	Parameters	Estimates	Std.Errors
	Linear Demand Estimates		
STAR 2	$\beta^{(STAR2)}$	-0.00	(0.023)
STAR 3	$\beta^{(STAR3)}$	-0.25	(0.037)
STAR 4	$\beta^{(STAR4)}$	-0.09	(0.043)
	Non Linear Demand Estimates		
Constant Clg.	$\beta_H^{(0)}$	7.67	(0.176)
Quality Clg	$\beta_H^{(q)}$	0.48	(0.041)
Price Intercept No Clg.	α_1^L	-0.04	(0.001)
Price Intercept Clg.	α_1^H	-0.06	(0.001)
Price Inc No Clg.	α_2^L	-0.01	(0.239)
Price Inc Clg.	α_2^H	-28.32	(1.157)
Travel Time L	λ^L	-0.19	(0.002)
Travel Time H	λ^H	-0.14	(0.006)

Notes: Estimates of the demand side coefficients. The specification also includes year and preschool fixed effects. Prices are in dollars per day and travel time in minutes.

Table 5: Selected Marginal Cost Estimates

Coefficients	Estimates	Std.Errors
Star 2	-0.14	(0.168)
Star 3	4.02	(1.116)
Star 4	5.96	(1.152)
Log 1 + Age	-3.01	(0.356)
Fraction Infant + Toddler	2.19	(0.8)
Log Density	1.31	(0.625)

Notes: Estimates of selected coefficients for centers' marginal costs. The specification also includes center and year fixed effects, and interaction with neighborhood demographics

6.2 Dynamic Parameters

I estimate CCP in a first-stage using a flexible multinomial logit. This estimates are used as a starting point for the PPI procedure. Following Sweeting (2013), in estimation I keep the transition kernel based on first-stage CCP for other players, implying that providers best-respond to beliefs about their competitors computed using first-stage CCP. This assumption is used to alleviate the computational burden and is relaxed when performing counterfactuals. In principle, the transition matrix $W(;\boldsymbol{\psi}^W)$ could include parameters for each authorized pairwise state transitions, but it is not possible to jointly identify scrap values, fixed costs, and entry costs in these models (Aguirregabiria and Suzuki, 2014). I fix both the scrap values and the transition to a lower quality rating to 0. I assume a common fixed costs for all providers, and allow for a fixed cost premium for centers operating above STAR 1.[38] The discount factor is chosen to be $\beta = 0.9$.

Table 6 presents the results of the PPI algorithm for estimation applied to the specification described above. First, the ranking of the transition costs are coherent: skipping a quality rating is rare in the data, and is estimated as being very costly. Second, the magnitude of the main fixed cost coefficient is plausible. If we consider an average STAR 4 in 2018, serving 45 children full time as suggested by Table 1, and consider a marginal cost of \$35 as suggested by Figure 6, this implies the following yearly cost per child[39]

$$\text{Yearly Cost Per Child STAR 4} = \frac{45 * 35 * 180 + (0.6 * 180 * 1000)}{45} = \$8,700$$

This estimate only reflects the flow cost, and does not include any of spending made to reach the

[38] As shown in Appendix F, almost all of the Exit happens at STAR 1. The fixed costs is estimated from the relative probability of staying in rather than exiting, which explains why this parameter is likely to pushed towards 0 in estimation.

[39] A year is taken as 180 days in the model.

highest-quality. It is close to industry estimates: slightly above the yearly costs of a public program in Pennsylvania operating centers rated STAR 3 & 4, evaluated at \$7,876 per child, and slightly below the \$9,543 spent per preschooler by Head Start in Pennsylvania (Friedman-Krauss et al., 2018). The entry parameter translates to a cost of \$313,000 to open a new center.[40] The standard deviation for private payoff, estimated at $0.47 * 180 * 1000 = \$84,600$ shocks is of the order of magnitude of variable profits recovered by the model. Table 9 in appendix F shows that the model fits the main moments of interest well. Upgrade rates are matched exactly, and while exit rates are slightly underestimated for higher quality firms, the model does a good job at matching exit by STAR 1 firms, which make up the bulk of the exiting providers.

Table 6: Dynamic Parameters

Coefficients	Estimates	Std. Err.
	Fixed Costs	
Fixed Cost	0.60	0.049
Add. Fixed Cost STAR 2,3,4	0.00	0.055
	Transition Costs	
Entry Cost	1.74	0.065
Upgrade 12	2.13	0.096
Upgrade 13	4.35	0.216
Upgrade 14	4.45	0.171
Upgrade 23	1.74	0.082
Upgrade 24	2.94	0.135
Upgrade 34	1.41	0.086
Upgrade Prev	0.31	0.067
Upgrade 12 x ξ	-0.05	0.031
Upgrade 23 x ξ	-0.03	0.042
	Private Payoff Shocks	
Std. Dev.	0.56	0.02
	Statistics	
Rsq. VF. Approx.	0.88	-
Neg. Log Likelihood	0.69	-

Notes: Results from the Nested Pseudo Likelihood estimator. Estimation is performed on a subset comprising most of the medium sized markets over the years 2010-2017. Value function approximation is performed using 2 duplicates of each observed state in the data with perturbed market structure and demographics. Coefficients levels can be converted into U.S. dollars by multiplying by 1000 times 180 days. Standard errors are not yet bootstrapped and only incorporate the inner-loop variation.

[40]For comparison, following the American Rescue Plan New York State awarded between \$398,000 and half a million dollars to grantees to open a center.(NY Division of Child Care Services, 2022).

7 Counterfactual Early Education Policies

7.1 Description of Policies

I use the estimated model to simulate counterfactual policy environments. The model endogenizes the main outcomes of interest for the design of early education policy: affordability, access, and quality. I rely on this richness and draw from recent policy proposals to simulate counterfactuals that can provide guidance on the design of ECE policies. Second, I use the counterfactuals to shed light on the main forces driving equilibrium responses to early education policies. The presence of both an extensive margin of response through providers' entry and exit and of an intensive margin through investments implies that policy makers face a trade-off between quantity and quality, as new entrants may divert enrollment from existing centers. In addition, market power can exacerbate the trade-off between efficiency and equity as centers' static price adjustments may raise costs for non-subsidized parents and decrease the quality attended by their children. To illustrate these trade-offs, I first design counterfactuals that work through static demand or supply channels, by directly targeting costs faced by parents for preschool or centers' revenue. I also simulate policies which directly create incentives for firms to enter, thereby changing families' access to quality. I present the simulated policy environments in the paragraphs below:

Pennsylvania ECE policies. This simulated environment corresponds to the policies in place in the data that was used in estimation. This environment, which captures the main features of the Pennsylvania Quality Rating and Improvement System (QRIS), has two components: demand subsidies to low income parents (below 200% FPL) and add-on rates compensating providers for serving subsidized children at high-quality.[41]

Increase in Add-on Rates. In the first set of expansion policies, I evaluate the impact of financial support targeted to providers serving subsidized children at high-quality. This counterfactual is inspired by add-on rates, a policy frequently used in QRIS, and is therefore labeled "Increased add-on".[42] I implement three levels of financial support: amounts equal to the 2018 maximum add-on rates (the highest level observed in the data) and levels of support two and three times as high. This

[41] See section 2 for a detail coverage of the Pennsylvania incentives scheme.

[42] Add-on rates, also called tiered-reimbursements, are a very common feature of QRIS: 18 out of the 26 state QRIS reviewed by Tout et al. (2010) feature a version of these incentives. There are two main types of tiered-reimbursements: fixed (dollar amount above market price) and variable (percentage of market price). In the counterfactuals as in estimation, I do not cap add-on rates to market prices; the increase in add-on is implemented as a per subsidized child transfer which only depends on the quality rating at which a center operates.

set of policies increases the variable profits of centers serving low-income children at high-quality. Depending on the strength of their effects on centers' variable profits, these policies can also have dynamic consequences on providers entry and upgrade decisions. For this set of counterfactuals, the demand-side subsidies stay the same as in the data.

Demand Subsidies Expansion. I evaluate a second approach to ECE expansion which focuses on lowering the costs faced by low and middle-income families. This policy environment maintains add-on rates to their levels in the data and (1) raises the eligibility threshold for subsides to 400% of the FPL, and by (2) lowers the co-payments to 7% of family income. This copayment schedule reflects what is considered an affordable ECE spending target in policy discussions, including in the Build Back Better agenda.[43]

Start-up Grants for Providers. I evaluate a third approach to ECE expansion which consists in offering start-up grant to opening early education providers. The amount of the grant is set equal to the estimated entry cost.[44] Instead of working through static channels like in previous environments, this policy proposal directly targets centers' incentives to enter. This policy, whose main goal is to increase geographic access to early education, can contribute to market expansion if families value proximity. Policies facilitating the entry of new providers have been enacted in the past in Florida, Indiana, Pennsylvania (Tout et al., 2010), and more recently New York (NY Division of Child Care Services, 2022). For this counterfactual, the demand-side subsidies stay the same as in the data.

7.2 Evaluation Criteria

A standard concern in the design of policies in education markets is that the planner's objective of improving students' outcomes is mediated through parents' choices. Parents may value attributes of education providers unrelated to skill formation, such as affordability and convenience, in particular in early education (Bassok et al., 2018). This feature of education markets favors a positive

[43]This target is an affordability target for child care in general, not specific to preschool. The BBB agenda discusses the implementation of universal preK for 3 and 4 year olds, an environment which is not simulated in the current version of this paper. Department of the Treasury (2021) discusses the administration's plan as "*a proposition to increase funding in the sector [...] and providing access to high-quality child care for low- and middle-income children. This child care plan will cut spending in half for most American families so that families do not have to spend more than 7 percent of their income on child care for young children*"

[44]The estimate for entry cost is $\$313,200$. An example of a similar policy recently implemented in the U.S is New York State's *Child Care Deserts Grant*. These opening grants ranged from $\$398,000$ to $\$500,000$ for child care centers, close to the amount of the proposed policy.

interpretation of the parameters of the demand system. In the counterfactual analysis, I primarily use the demand system to compute equilibrium enrollment in alternative policy environments. Even though I compute consumer surplus to get a sense of the changes in attributes valued by parents implied by the policies, I do not use it to rank the policies. In a similar spirit as recent structural papers studying education markets (Neilson, 2013; Allende, 2019; Armona and Cao, 2022; Bates et al., 2022), I evaluate policies on their educational effectiveness, and I take the Keystone Star rating as a measure of a provider's educational productivity. The underlying assumption in the following analysis is that, from the perspective of educational effectiveness, a child enrolled in high-quality is preferred to a child enrolled in low-quality, which is itself preferred to the outside option.[45]

For each new policy environment, I simulate the equilibrium path under the counterfactual and compare it to a baseline environment where the demand-subsidies and add-on rates in place in the data are removed. This baseline therefore represents and environment where outcomes are primarily determined by parents preferences and centers' costs. The main market outcomes of interests are the quality of the supply of centers and the distribution of enrollment by quality ratings over the income distribution of parents. In addition, I compute the change in consumer surplus by income. For a given policy environment Υ in year t, for a demographic group b, this measure can be computed from the estimated demand model by aggregating expected utilities within group b and normalizing by individual price coefficients:

$$CS_t^b(\Upsilon) = \sum_{i \in b} \frac{1}{|\alpha_i|} \frac{w_i}{\sum_{i \in b} w_i} \log(1 + \sum_{\mathcal{J}_t} \exp(u_{ijt}))$$

Lastly, I compute the cost associated with each policy environment, which enables me to compare the effectiveness of policies at achieving a given objective. If the policy maker's goal is to raise enrollment (in high-quality centers), the effectiveness can be defined as the additional (high-quality) enrollment per dollar spent.[46] Designing policies that are ex-ante budget neutral is challenging in this context, because costs depend on the endogenous take-up of policies from both parents and providers. Therefore, I instead compare policies by scaling their ex-post average costs by outcomes of interest, such as additional enrollment or centers in operation.

[45]The assumption that a child enrolled in the formal ECE market is better-off than in informal options follows from the papers studying the effect of ECE arrangements on child's outcomes. This assumption may be more debatable for high-income parents where the outside option may involve high-quality tutoring of parental time.

[46]This is similar to Dunne et al. (2013)'s paper on entry subsidies where policies are evaluated in terms dollar per additional firm in operation.

7.3 Solution Method For Counterfactuals

The simulated policies may change the mapping between states and centers' revenues which implies that centers' strategies computed during estimation are not an equilibrium in the new environment. Each counterfactual therefore requires solving for a new set of centers' strategies. While there is no need to search over dynamic parameters, the difficulties associated with finding a solution to the dynamic game outlined in estimation are still a challenge for counterfactual simulations.[47] The relevant state space for firms' decisions is high-dimensional, making the computation of value functions intractable. Therefore, the solution method for counterfactuals also relies on parametric policy iteration and approximation of the value function. As in estimation, I use neural networks to reliably predict variable profits in many configurations of the state and construct the approximation basis for the value function. Because the mapping of states to preschools' variable profits changes across policy environments, each counterfactual requires training a new neural network.

The counterfactual under a new policy environment Υ starts with centers placed in the same states as in the data in 2010. While centers' states are the same, variable profits may differ due to the change in policy. I then compute a new vector of equilibrium strategies $\mathbb{P}^*_{2010}(\Upsilon)$, and simulate the market structure in 2011 by drawing centers' states in 2011 according to $\mathbb{P}^*_{2010}(\Upsilon)$. The remaining stochastic components of the 2011 market state $\mathcal{M}_{2011}$ are drawn according to their estimated distributions. This process is then repeated sequentially to recover the path of centers' states over time. The simulations run until 2017 in order to evaluate the long-run effect of the policy on market structure and enrollment. Given the path of market states $\{\mathcal{M}_t(\Upsilon)\}_{t=2010,\ldots,2017}$, I compute the static equilibrium in each market state to recover the outcomes of interest.

To compute equilibrium strategies $\mathbb{P}^*_t(\Upsilon)$, I use the same solution approach as in estimation and discretize the state space to perform value function approximation. For each counterfactual, I form a new basis of N points, compute variable profits under the policy environment Υ on this grid and train a neural network to predict $\pi(\mathcal{M}_{xjt};\Upsilon)$. The predicted profits are then used to construct approximation bases relevant for the value function. I provide additional details on the solution method in Appendix G.

[47] An additional difficulty in counterfactual simulations is the possibility of multiple equilibria. In estimation this issue was avoided by fixing beliefs about competitors to first-stage CCPs, but these CCPs would not provide a plausible approximation of firms' beliefs in a new policy environment. To solve for counterfactuals, I use iterated best-response to update both startegies and beliefs about competitors' actions. This approach could in principle encounter multiple equilibria, but I have not faced this issue in the counterfactuals presented above. The estimated variance of private payoff shocks ψ^ε is large relative to variable profits, making this problem less acute (Sweeting, 2013)

7.4 Results from Policy Simulations

The dynamic effect of early education policies. Figure 7 shows the simulated of the supply of each quality rating under the different policy environments. Two general patterns stand out from these trajectories. First, the market structure takes time to adjust, which highlights the importance of considering both the short but also the long-run impact policies can have on ECE access and quality. Second, the changes relative to the baseline environment suggest increases in the number of centers, consistent with policies aiming at expanding the ECE market.

Policies in the Pennsylvania environment increase the supply of higher quality centers, particularly at the medium rating STAR 2: the top left panel shows that in 2017, the number of STAR 2 centers in operation is more 50% higher that in the baseline. Compared to this environment, increasing the add-on rate has a large effect on the supply of the highest quality centers, for which the incentives created by the add-on rate are the strongest. The top center panel shows that had the add-on rate been at its highest level ($9.20 for STAR 4) for the entire period 2010-2018, the supply of STAR 3 & 4 centers would have increased by almost 100%. This effect jumps to 300% if the add-on rate is increased threefold. The already large gap in high-quality supply between the PA environment, which features small add-on levels, and the Add-On at its 2018 level suggests that the levels of the policy used for most of the sample in the data were not large enough to induce centers to invest in high-quality provision. This is consistent with the estimated large upgrade and marginal costs associated with higher ratings, together with relatively small willingness to pay from parents. This set of policy environments highlights the potential of static policies to change centers' dynamic decisions through their impact on revenue and influence market structure over time.

The demand expansion shown on the bottom center panel of Figure 7 produces an increase in the supply of high-quality centers similar in magnitude to the small increase in add-on environment. Raising the income eligibility threshold to cover middle-income families generates incentives for providers to upgrade through two channels. First, the policy increases the demand from college-educated, quality-sensitive parents. Provided that families do not live in fully segregated neighborhoods, centers serving children from mixed SES invest in quality to cater to the demand of high and middle-income families nearby, generating a form of preference externalities from the perspective of the planner. Second, by expanding the set of subsidized families, this policy increases the total amount of financial support centers receive from add-on rates, as these rates now apply to a broader set of subsidized children (see the small increase in add-on expenses in Table 7).

The bottom right panel of Figure 7 displays the market structure under the environment with start-up grants. This additional policy generates entry which translates into an increasing stock of STAR 1, and STAR 2 centers over the years. However, the number of high-quality centers barely increases, implying lower rates of upgrade than in the other environments. These lower upgrading rates reflect a business stealing effect due to the large number of STAR 1 centers poaching students away from competitors, thereby decreasing profits and incentives to invest in high-quality. This effect is not surprising given the high willingness to avoid travel recovered in estimation. This business stealing channel reinforces the quantity-quality trade-off faced by the policy maker: additional quantity may be provided at the expense of future quality due to lower revenues for competing providers.

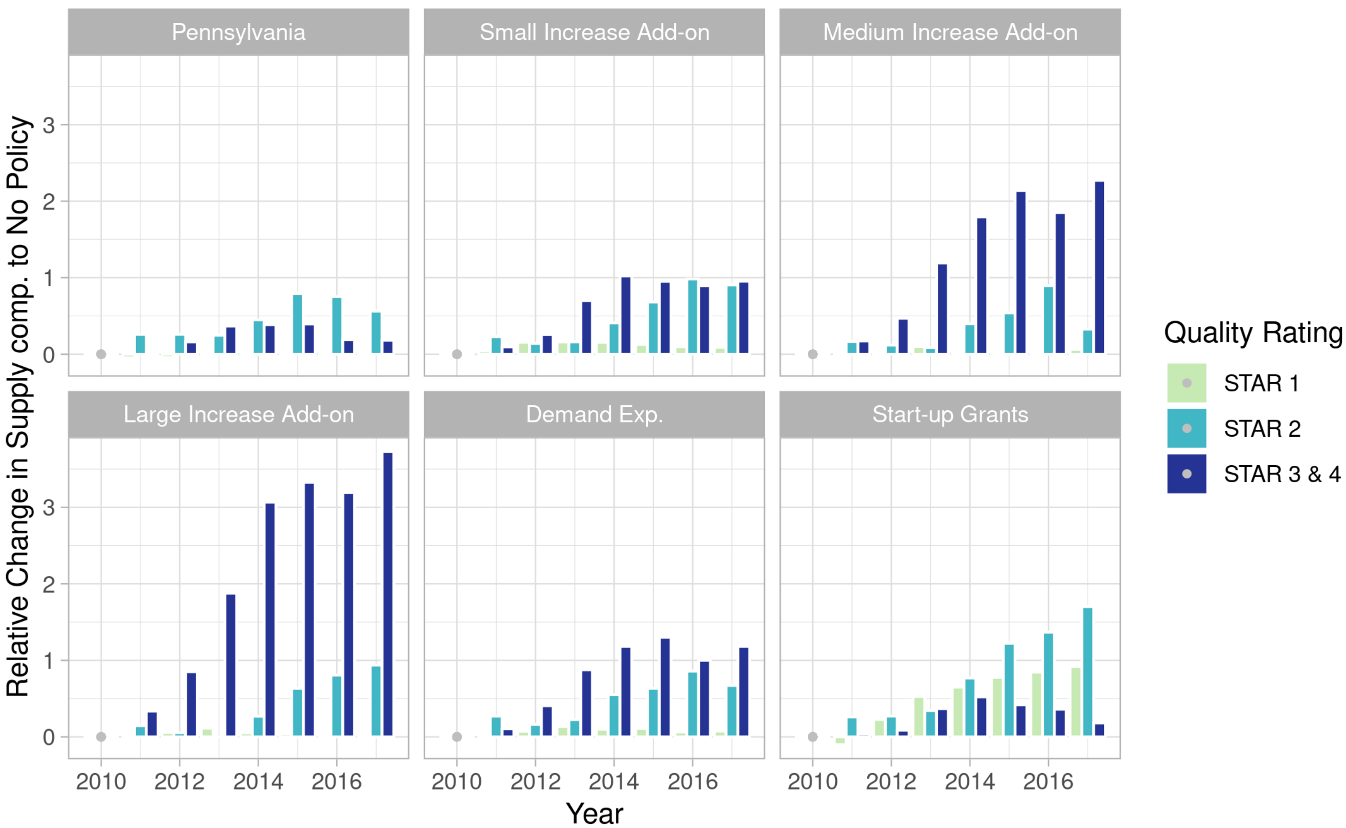


Figure 7: Market Structure Paths in Selected Counterfactuals

Notes: The figure shows the relative change in the number of centers in each STAR rating, in each year, for each counterfactual, compared to the baseline environment. 2010 corresponds to no change as all the environments are initialized in the same state as the one observed in the data. PA stands for the environment with the same policies as in the data.

Cost-effectiveness of Simulated Policies. To compare the performance of these policies, I compute the average cost associated with each environment over the simulation path. I also calculate the additional number of centers and (high-quality) enrollment they generate and form 3 measures of cost-effectiveness: dollar per preschooler, per preschooler in high-quality center, and per center, all relative to the baseline environment. These measures of are reported in the *Cost per Outcome* panel of Table 7 and can be used to rank the policies according to their aggregate effectiveness. These measures suggest that the increases in add-on rates perform well along all effectiveness measures. The small increase in add-on rates is most effective at market expansion, with only $8,093 per additional preschooler. The large increase in add-on rate is the most effective at high-quality expansion, with $7,731 per additional preschooler in high quality.[48] The *Outcomes* columns of Table 7 allow to decompose the effects of the policy. The large add-on increase achieves this level of effectiveness in part through market expansion (almost 4,000 additional children enrolled) but also through reallocation of children enrolled to higher quality centers (the number of additional children in high-quality is higher than the number of additional children). Finally, add-on policies almost match start-up grants in their effectiveness at raising the supply of centers: add-on policies and start-up grants all cost around $300,000 per additional center. While start-up grants generate a lot of entry, many centers still exit the market. In contrast, increase in add-on rates both induce new supply and stabilize existing centers by raising revenues, and achieve this at a much lower total cost.

The shortcomings of the demand expansion and start-up grants provide insights into parents and centers' equilibrium responses to policies. Demand expansion to middle income parents increases enrollment by 5,048, more than any other policy, but does so at a cost of more than $66 million. The high cost of this policy is in part due to centers' price responses which reduce the pass-through to families and increase the cost of the demand subsidies. The large impact of this policy on market expansion nevertheless suggests that cost is an important barrier for middle-income parents. Some degree of demand expansion with a less generous co-payment could be used in combination with incentives for quality provision to design a more effective policy. Finally, start-up grants are not a very cost-effective policy to generate high-quality enrollment. While this policy does generate market expansion (4,098 new preschoolers), it mostly displaces enrollment to low-quality preschools. This suggests that families' preferences for proximity create a sizable business stealing effect, whereby families substitute away from high-quality centers to the newly open, convenient but low-quality options.

[48] This cost per child is in line with estimates of typical preschool programs, and is even quite low for a program running at high-quality (Friedman-Krauss et al., 2018). But here a large fraction of the costs of the systems are still borne by middle and high-income parents, who fully pay for their enrollment.

Table 7: Cost, Outcomes and Effectiveness Measures of Simulated Policies

	Cost (million USD)				Outcomes		Cost per Outcome		
Environment	Total	Add-on	Demand Sub.	Dynamic Incentive	Extra Enrollment	Extra High-quality Enrollment	Dollar per Child	Dollar per Child in High Quality	Dollar per Center
Pennsylvania	26.60	0.66	25.94	0.00	3223.00	1186.00	8253.00	22420.00	816713.00
Small Increase Add-on	29.26	2.12	27.13	0.00	3615.00	1924.00	8093.00	15203.00	302969.00
Medium Increase Add-on	32.99	5.88	27.11	0.00	3680.00	3403.00	8964.00	9693.00	382950.00
Large Increase Add-on	39.02	11.40	27.62	0.00	3947.00	5048.00	9886.00	7731.00	310064.00
Demand Exp.	66.29	1.65	64.65	0.00	5307.00	2840.00	12492.00	23346.00	723935.00
Start-up Grants	68.16	0.61	28.52	39.02	4098.00	814.00	16633.00	83718.00	297998.00

Notes: Values in the table are computed relative to the baseline environment. I take the average over 2010-2016 of costs and outcomes, and compute the inverse effectiveness measures as the ratios of average costs to average outcomes. Dynamic Incentive stands for cost of policies directly targeting entry or investment. High-quality stands for STAR 3 and STAR 4 centers. All outcomes are in counts, and inverse-effectiveness metrics in dollars spent per extra preschooler/center.

Table 8: Policy Consequences on Enrollment and Consumer Surplus by Income

		Less than 2*FPL		[2*FPL, 4*FPL)		More than 4*FPL	
Environment	Total Cost (million USD)	Consumer Surplus	Share Enrolled	Consumer Surplus	Share Enrolled	Consumer Surplus	Share Enrolled
Pennsylvania	26.60	1.56	1.06	-0.08	-0.01	-0.02	0.01
Small Increase Add-on	29.26	1.83	1.19	-0.03	0.02	0.02	0.02
Medium Increase Add-on	32.99	1.90	1.22	0.00	0.03	0.05	0.03
Large Increase Add-on	39.02	2.07	1.29	0.05	0.06	0.08	0.04
Demand Exp.	66.29	2.24	1.31	0.99	0.41	-0.13	-0.04
Start-up Grants	68.16	2.30	1.43	0.07	0.09	0.10	0.06

Notes: Values in the table are computed relative to the baseline environment. Table reports the percentage change in consumer surplus and enrollment for each policy environment. Cost is computed as the average over the path, while changes in enrollment and consumer surplus are computed in 2016 to represent the long-run effect of policies. Families are grouped in 3 bins by income based constructed from the Federal Poverty Line (FPL). Households below 200% of the FPL is the group subsidized in the data (Pennsylvania environment).

Distributional Consequences. Next, I use the demand model to break down outcomes by household income and analyze the distributional consequences of the policies. From the perspective of a planner who values educational efficiency and equity, enrollment by income and quality are the relevant outcomes to evaluate the policies. Figure 8 shows the percentage point change in enrollment relative to baseline by STAR rating and parental by income. First, the distribution of aggregate enrollment changes highlight the importance of demand subsidies in determining the degree of market expansion. In the policy increasing add-on (left panel) only low-income parents receive subsidies, and market expansion is entirely driven by this group, whose enrollment increases by almost 15 p.p across all ratings. Similarly in the middle panel, where both low and middle income parents are subsidized, market expansion is driven by these two income groups. In the middle panel, aggregate enrollment of high-income children even decreases, due to centers' exerting market power to raise prices in response to the demand subsidies. However, looking at the breakdown of enrollment by STAR ratings shows that in these two policies there is a reallocation of children from low-quality to high-quality centers, and this reallocation happens even for non-targeted groups. For instance, enrollment of high-income children in STAR 3&4 increases by more than 6 p.p after the demand expansion policy. If policies targeted to a sub group of parents generate incentives for providers to upgrade, they may induce quality spillovers for non-targeted parents. Providers dynamic responses therefore mitigate the equity efficiency trade-off from the perspective of the planner.

Consumer surplus illustrates the dimensions along which parents preferences and the planner's objective of increasing educational efficiency may differ. Table 8 reports the distribution of enrollment change and consumer surplus by income group. The start-up grant scenario results in the larger increases in consumer surplus for all income groups (230%, 7%, and 10%) compared to the large in add-on policy (207%, 5%, and 8%), even though the latter is much more effective at increasing the quality of the supply. This is due to the fact that start-up grants increase the variety of nearby options for families. Middle-income families experience higher consumer surplus in the demand expansion scenario even though their enrollment in high-quality is higher under the high add-on rates regime. This tension between parents' value for convenience and the objective to increase high-quality enrollment is characteristic of the early education market, and is for instance discussed in Herbst et al. (2020).

Counterfactual Access to high-quality ECE. Finally, I consider the influence of each policy environment on the spatial distribution of ECE. Figure 9 shows centers in activity and their quality

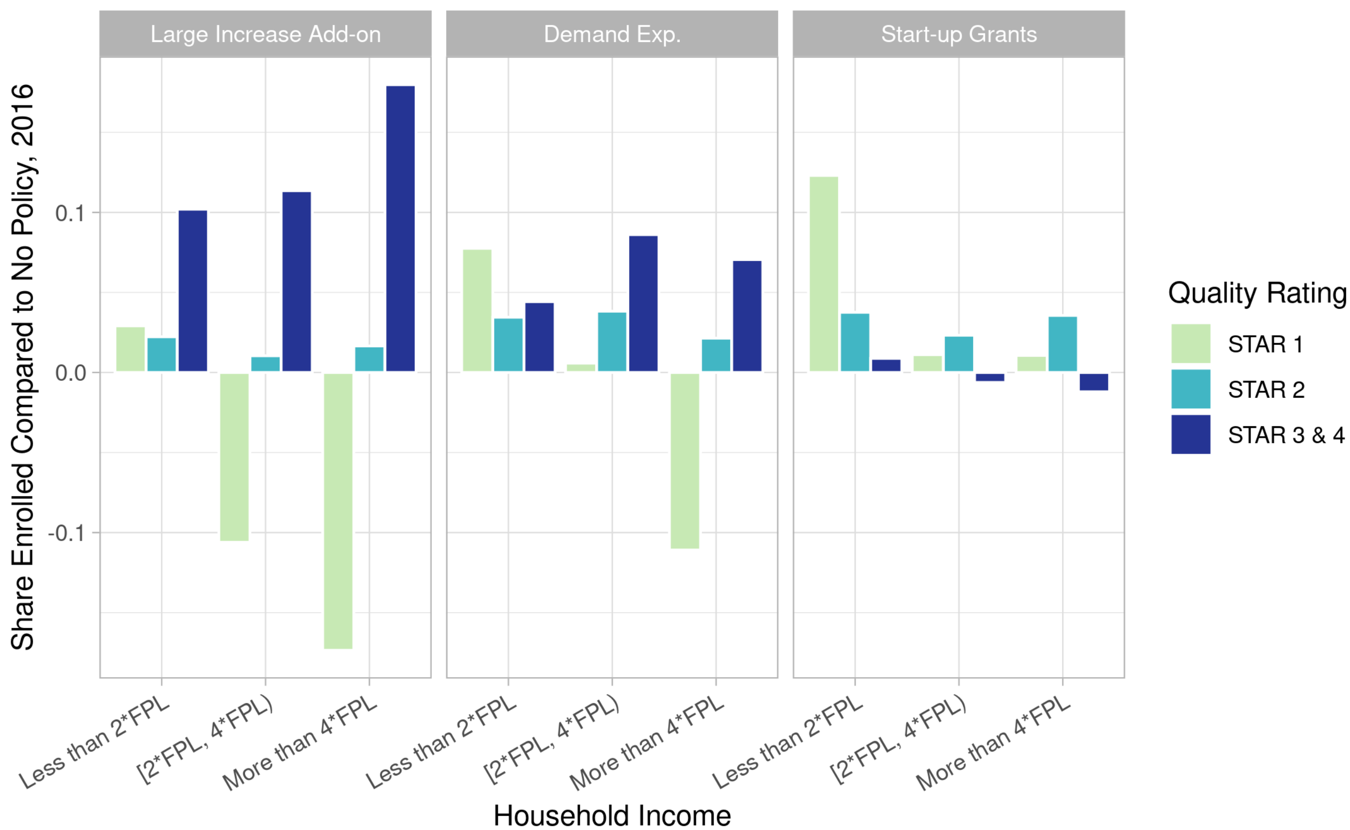
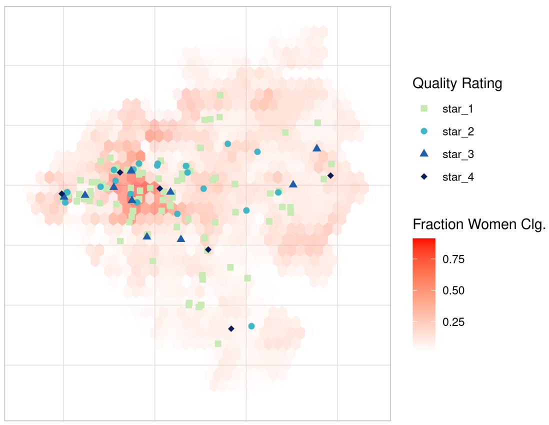


Figure 8: Distributional Consequences: Enrollment by Star Rating and Income

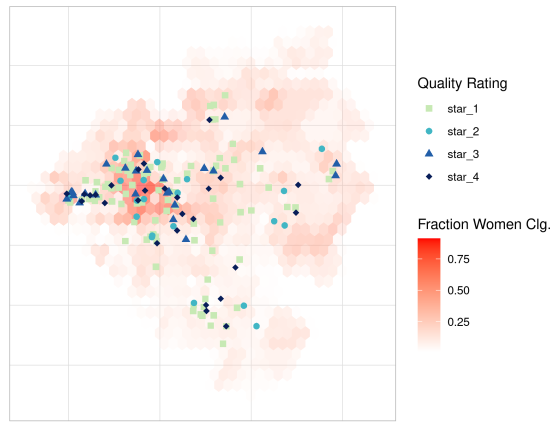
Notes: Percentage Point change in enrollment, broken down by Star rating and income bins of families.

ratings superimposed on the fraction of college graduate women across East Pittsburgh in 2016. The Figure highlights how the forces described in the previous paragraphs take shape in space. As can be seen on the bottom-right panel, start-up grants lead to an increase in the number of centers at the lowest rating across the market. The ubiquity of competition makes it difficult for providers to increase profits by attracting additional children, as parents have a large set of convenient options around. This results in an environment with very few high-quality providers. The difference between the distribution of quality in the large increase in add-on compared to the demand expansion highlights the role of add-on rates in relaxing the spatial matching between quality and demographics. The demand expansion (bottom left) results in an increase in demand for quality sensitive parents which leads to more centers upgrading to higher ratings. However, these upgrades occur primarily in the affluent, highly-educated areas of Pittsburgh identified in Section 3, in the darker red regions near the western tip of the map. In contrast, add-on rates are a form of doubly targeted policies, making public funding conditional on serving low-income children at high-quality. The large increase in add-on rates results in centers upgrading to the highest ratings not only in the affluent neighborhoods, but also on their periphery and at the southern tip of the market, as can be seen on the top right panel of Figure 9. Even though targeted to firms serving low-income children, the map also makes clear why this policy also benefits high and middle class families, as shown in the consumer surplus measures of Table 8. Neighborhoods are not completely segregated by income, and incentives created by targeted add-on rates also increase the availability of high-quality options for medium and high-income parents who value higher ratings.

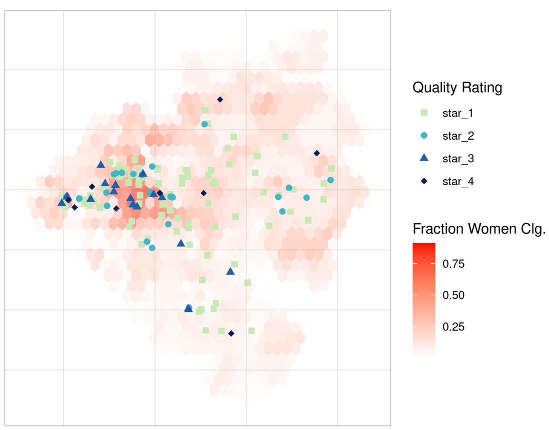
Summary of the results. Three main takeaways from these results can inform the design of ECE policies. First, the dynamic responses of providers are key to evaluate the impact of policies. Access and quality take time to adjust as providers enter, exit and upgrade over the years. Policies directed at static variables also influence dynamic decisions: raising the profits of providers at a given quality-rating in specific locations eventually changes the supply in these areas through entries and upgrades. Second, market power amplifies the trade-off between equity and efficiency: in the case of broad demand subsidies to low and middle-income families, high-income parents suffer a 12% reduction in welfare due to price responses. Policies that mitigate these price adjustments, such as add-on rates which substitute public funds to firms' markups, do not harm non-targeted parents. These results emphasize the importance of taking into account both providers' static price adjustments and dynamic responses when designing early education policies. Lastly, the counterfactuals shed



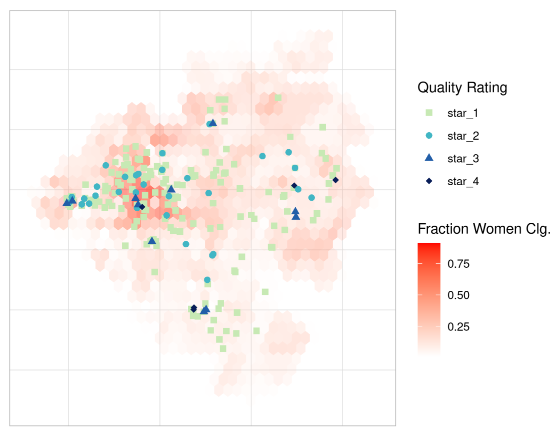
(a) Pennsylvania



(b) Large Increase in Add-on



(c) Demand Expansion



(d) Start-up Grants

Figure 9: Counterfactual Geographic Distribution and Quality of ECE in Pittsburgh

Notes: The maps display the centers in operation and their ratings in each selected policy environments in 2016.

light on both the type of market-based approaches and the levels of policies that can address the main challenges of designing policies in ECE markets. The difficulty of increasing enrollment of low-income children in high-quality centers stems from parents' value for proximity, competition between local providers with market power, and high fixed and operating costs associated with high-quality provision. I find that a policy offering financial support targeted to high-quality centers serving low-income children, similar to the add-on component of the Pennsylvania QRIS but at higher rates, would lead to more low and middle-income children enrolled in high-quality at a lower cost than both an expansion of demand subsidies or generous start-up grants. Moreover, the counterfactual simulations show how the effectiveness of this policy varies with its generosity. As the level of financial support increases and gradually overcomes providers' upgrading costs, I find that not only high-quality enrollment increases, but also that the cost per preschooler in high-quality falls. This policy could be combined with quality-targeted policies on the demand-side as well.

8 Conclusion

The potential of high-quality preschool to reduce educational inequality before kindergarten has created policy interest in increasing public spending in early education. The design of effective policies requires understanding private providers' equilibrium responses. Providers tend to be small firms, many of them operating at low-quality. Payments mostly come from parents and many children are not served at all. The success of a policy at increasing high-quality enrollment for low-income children therefore depends both on providers price responses and on their decisions to enter and invest in quality in targeted neighborhoods.

In this paper, I develop a dynamic equilibrium model of the preschool market to provide guidance on the design of early education policies. I first use the model to highlight the challenges of increasing high-quality enrollment for disadvantaged children. Preschools have market power, parents trade-off quality with affordability and proximity, and providing high-quality is costly. Second, I evaluate the aggregate and distributional effects of counterfactual early education policies in equilibrium. Policies which tie financial support to high-quality provision and serving low-income children are more cost-effective than proposed alternatives at increasing high-quality enrollment for children from all income levels. This result is informative for the design of successful QRIS, an approach to early education that most U.S states are currently implementing or developing.

In this analysis of the early education market, I do not investigate the inputs in the production of high-quality preschool. Avenues for future research include linking the study of ECE providers with the preschool teachers' labor market, as the training, hiring, and retention of educators are important aspects of ECE provision.

References

Aguirregabiria, V. and P. Mira (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75, 1–53.

——— (2010): "Stability of Equilibria and Properties of Sequential Estimators: Problems and Solutions," *Manuscript. University of Toronto.*

Aguirregabiria, V. and J. Suzuki (2014): "Identification and counterfactuals in dynamic models of market entry and exit," *Quantitative Marketing and Economics*, 12, 267–304.

Allende, C. (2019): "Competition Under Social Interactions and the Design of Education Policies," *Job Market Paper.*

Anderson, D. G. (1965): "Iterative procedures for nonlinear integral equations," *Journal of the ACM (JACM)*, 12, 547–560.

Armona, L. and S. Cao (2022): "Redesigning Federal Student Aid in Higher Education," .

Backus, M., C. Conlon, and M. Sinkinson (2021): "Common ownership and competition in the ready-to-eat cereal industry," Tech. rep., National Bureau of Economic Research.

Baker, M., J. Gruber, and K. Milligan (2008): "Universal child care, maternal labor supply, and family well-being," *Journal of Political Economy*, 116, 709–745.

——— (2019): "The long-run impacts of a universal child care program," *American Economic Journal: Economic Policy*, 11, 1–26.

Bassok, D., M. Fitzpatrick, and S. Loeb (2011): "Disparities in Child Care Availability across Communities: Differential reflection of targeted interventions and local demand," *Center for Education Policy Analysis.*

——— (2014): "Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia," *Journal of Urban Economics*, 83, 18–33.

Bassok, D. and E. Galdo (2016): "Inequality in preschool quality? Community-level disparities in access to high-quality learning environments," *Early Education and Development*, 27, 128–144.

Bassok, D., P. Magouirk, A. J. Markowitz, and D. Player (2018): "Are there differences in parents' preferences and search processes across preschool types? Evidence from Louisiana," *Early Childhood Research Quarterly*, 44, 43–54.

Bates, M. D., M. Dinerstein, A. C. Johnston, and I. Sorkin (2022): "Teacher Labor Market Equilibrium and Student Achievement," Tech. rep., National Bureau of Economic Research.

Beatty, B. (1995): *Preschool education in America: The culture of young children from the colonial era to the present*, Yale University Press.

Beresteanu, A. and P. B. Ellickson (2019): "The Dynamics of Retail Oligopoly," .

Berlinski, S., M. M. Ferreyra, L. Flabbi, and J. D. Martin (2020): "Child care markets, Parental labor supply, and child development," *Parental Labor Supply, and Child Development*.

Bernal, R. and M. P. Keane (2011): "Child care choices and children's cognitive achievement: The case of single mothers," *Journal of Labor Economics*, 29, 459–512.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 841–890.

——— (1999): "Voluntary export restraints on automobiles: Evaluating a trade policy," *American Economic Review*, 89, 400–430.

Blau, D. and J. Currie (2006): "Pre-school, day care, and after-school care: who's minding the kids?" *Handbook of the Economics of Education*, 2, 1163–1278.

Borowsky, J., J. H. Brown, E. E. Davis, C. Gibbs, C. M. Herbst, A. Sojourner, E. Tekin, and M. J. Wiswall (2022): "An equilibrium model of the impact of increased public investment in early childhood education," Tech. rep., National Bureau of Economic Research.

Brown, J. H. (2018): "Does public pre-k have unintended consequences on the child care market for infants and toddlers?" *Princeton University Industrial Relations Section Working Paper*, 626.

Bugni, F. A. and J. Bunting (2021): "On the iterated estimation of dynamic discrete choice games," *The Review of Economic Studies*, 88, 1031–1073.

BUILD Initiative (2017): "Finance and Quality Rating and Improvement Systems," .

Cahan, E. D. (1989): *Past caring: A history of US preschool care and education for the poor, 1820–1965.*, National Center for Children in Poverty.

Cannon, J. S., G. L. Zellman, L. A. Karoly, and H. L. Schwartz (2017): *Quality rating and improvement systems for early care and education programs: Making the second generation better*, RAND.

Caoui, E. H., B. Hollenbeck, and M. Osborne (2022): "The Impact of Dollar Store Expansion on Local Market Structure and Food Access," .

Cascio, E. U. (2015): "The promises and pitfalls of universal early education," *IZA World of Labor.*

Cascio, E. U. and D. W. Schanzenbach (2013): "The impacts of expanding access to high-quality preschool education," Tech. rep., National Bureau of Economic Research.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and Y. Feng (2019): "On binscatter," *arXiv preprint arXiv:1902.09608.*

Collard-Wexler, A. (2013): "Demand fluctuations in the ready-mix concrete industry," *Econometrica*, 81, 1003–1037.

Conlon, C. and J. Gortmaker (2020): "Best practices for differentiated products demand estimation with pyblp," *The RAND Journal of Economics*, 51, 1108–1161.

Cui, J. and R. Hanson (2021): "Early Child Care in Single-Parent and Two-Parent Families: 2019. Data Point. NCES 2021-005." *National Center for Education Statistics.*

Davis, E. E., W. F. Lee, and A. Sojourner (2019): "Family-centered measures of access to early care and education," *Early Childhood Research Quarterly*, 47, 472–486.

Davis, E. E. and A. Sojourner (2021): "Increasing Federal Investment in Children's Early Care and Education to Raise Quality, Access, and Affordability," *Washington. DC: Hamilton Project.*

Dearing, A. and J. R. Blevins (2019): "Efficient and convergent sequential pseudo-likelihood estimation of dynamic discrete games," *arXiv preprint arXiv:1912.10488.*

Department of the Treasury (2021): *The Economics of Child Care Supply in the United States*, United States Department of the Treasury.

Dinerstein, M., C. Neilson, and S. Otero (2020): *The Equilibrium Effects of Public Provision in Education Markets: Evidence from a Public School Expansion Policy*, Princeton University, Industrial Relations Section.

Dubé, J.-P., A. Hortaçsu, and J. Joo (2021): "Random-coefficients logit demand estimation with zero-valued market shares," *Marketing Science*, 40, 637–660.

Duncan, G. J. and A. J. Sojourner (2013): "Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps?" *Journal of human resources*, 48, 945–968.

Dunne, T., S. D. Klimek, M. J. Roberts, and D. Y. Xu (2013): "Entry, exit, and the determinants of market structure," *The RAND Journal of Economics*, 44, 462–487.

Durkin, K., M. W. Lipsey, D. C. Farran, and S. E. Wiesen (2022): "Effects of a statewide pre-kindergarten program on children's achievement and behavior through sixth grade." *Developmental Psychology*.

Elango, S., J. L. García, J. J. Heckman, and A. Hojman (2015): "Early childhood education," in *Economics of Means-Tested Transfer Programs in the United States, Volume 2*, University of Chicago Press, 235–297.

Elliott, J., G. V. Houngbonon, M. Ivaldi, P. T. Scott, et al. (2021): *Market structure, investment and technical efficiencies in mobile telecommunications*, Centre for Economic Policy Research.

Friedman-Krauss, A. H., K. S. Barnett, G. Weisenfeld, and N. DiCrecchio (2018): "The State of Preschool 2017: State Preschool Yearbook." *National Institute for Early Education Research*.

Gandhi, A. and J.-F. Houde (2019): "Measuring substitution patterns in differentiated-products industries," *NBER Working Paper*.

Gandhi, A., Z. Lu, and X. Shi (2020): "Estimating demand for differentiated products with zeroes in market share data," *Available at SSRN 3503565*.

Gowrisankaran, G., A. Langer, and W. Zhang (2022): "Policy uncertainty in the market for coal electricity: The case of air toxics standards," Tech. rep., National Bureau of Economic Research.

Gray-Lobe, G., P. A. Pathak, and C. R. Walters (2021): "The long-term effects of universal preschool in Boston," Tech. rep., National Bureau of Economic Research.

Grunewald, R. and P. Davis (2011): "Hardly child's play," *Federal Reserve Bank of Minneapolis.*

Heckman, J., R. Pinto, and P. Savelyev (2013): "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American Economic Review*, 103, 2052–86.

Hendren, N. and B. Sprung-Keyser (2020): "A unified welfare analysis of government policies," *The Quarterly Journal of Economics*, 135, 1209–1318.

Herbst, C. M. (2018): "The impact of quality rating and improvement systems on families' child care choices and the supply of child care labor," *Labour Economics*, 54, 172–190.

Herbst, C. M., K. C. Desouza, S. Al-Ashri, S. S. Kandala, M. Khullar, and V. Bajaj (2020): "What do parents value in a child care provider? Evidence from Yelp consumer reviews," *Early Childhood Research Quarterly*, 51, 288–306.

Hotz, V. J. and M. Wiswall (2019): "Child care and child care policy: existing policies, their effects, and reforms," *The ANNALS of the American Academy of Political and Social Science*, 686, 310–338.

Hotz, V. J. and M. Xiao (2011): "The impact of regulations on the supply and quality of care in child care markets," *American Economic Review*, 101, 1775–1805.

Kamerman, S. B. and S. Gatenio-Gabel (2007): "Early childhood education and care in the United States: An overview of the current policy picture," *International Journal of Child Care and Education Policy*, 1, 23–34.

Keys, T. D., G. Farkas, M. R. Burchinal, G. J. Duncan, D. L. Vandell, W. Li, E. A. Ruzek, and C. Howes (2013): "Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics," *Child development*, 84, 1171–1190.

Kline, P. and C. R. Walters (2016): "Evaluating public programs with close substitutes: The case of Head Start," *The Quarterly Journal of Economics*, 131, 1795–1848.

Ludwig, J. and D. L. Miller (2007): "Does Head Start improve children's life chances? Evidence from a regression discontinuity design," *The Quarterly journal of economics*, 122, 159–208.

McFarland, J., B. Hussar, J. Zhang, X. Wang, K. Wang, S. Hein, M. Diliberti, E. F. Cataldi, F. B. Mann, and A. Barmer (2019): "The Condition of Education 2019. NCES 2019-144." *National Center for Education Statistics.*

Morrow, W. R. and S. J. Skerlos (2011): "Fixed-point approaches to computing Bertrand-Nash equilibrium prices under mixed-logit demand," *Operations research*, 59, 328–345.

National Center for Education Statistics (2022): "Enrollment Rates of Young Children," *Condition of Education.*

Neilson, C. (2013): "Targeted vouchers, competition among schools, and the academic achievement of poor students," *Job Market Paper*, 1–62.

NSECE Project Team (2014): "Characteristics of Center-based Early Care and Education Programs: Initial Findings from the National Survey of Early Care and Education," .

NY Division of Child Care Services (2022): "Invest in New York - Child Care Deserts Grant for New Providers," https://ocfs.ny.gov/programs/childcare/deserts/#awardees, accessed: 2022-11-01.

OECD (2019): *Education at a Glance 2019.*

OPRE (2016): "Fact Sheet: How Far Are Early Care and Education Arrangements from Children's Homes?" *OPRE Report*, No. 2016-10.

Pesendorfer, M. and P. Schmidt-Dengler (2010): "Sequential estimation of dynamic discrete games: A comment," *Econometrica*, 78, 833–842.

Rosenthal, J. (1971): "PRESIDENT VETOES CHILD CARE PLAN AS IRRESPONSIBLE," *New York Times.*

Rust, J. (2000): "Parametric Policy Iteration: An Efficient Algorithm for Solving Multidimensional DP Problems?" Tech. rep., Mimeo, Yale University.

Ryan, S. P. (2012): "The costs of environmental regulation in a concentrated industry," *Econometrica*, 80, 1019–1061.

Seim, K. (2006): "An empirical model of firm entry with endogenous product-type choices," *The RAND Journal of Economics*, 37, 619–640.

Singleton, J. D. (2019): "Incentives and the supply of effective charter schools," *American Economic Review*, 109, 2568–2612.

Sun, L. and S. Abraham (2021): "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 225, 175–199.

Sweeting, A. (2013): "Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81, 1763–1803.

Tout, K., R. Starr, M. Soli, S. Moodie, G. Kirby, and K. Boller (2010): "Compendium of Quality Rating Systems and Evaluations: The Child Care Quality Rating System (QRS) Assessment." *Administration for Children & Families.*

U.S. Department of Education (2014): "U.S. Department of Education National Center for Education Statistics, Digest of Education Statistics," https://nces.ed.gov/programs/digest/d19/tables/dt19_202.60.asp, accessed: 2022-12-30.

A Notation

Symbol	Description
m	index of geographic market
t	index of school year (September yr - August $yr+1$)
j	index of preschools
e	index of mother's educational attainment
ℓ	index of neighborhood (block-group)
i	index of consumer
b	index of income bin
τ	index of subsidy status
n	index of income nodes from quadrature (consumer types)
$\mathcal{L}_m$	set of neighborhoods in market m
$\mathcal{J}_{mt}$	set of preschools operating in geographic market m at t
$\mathcal{J}_{mt}^{obs}$, $(\mathcal{J}_{mt}^{miss})$	preschools with observed (missing) market shares in market m at t
s_{jt} (s_{jt}^k)	aggregate (subsidy specific) market share of product j
δ_{jt}	product specific utility to consumers
$\mu_{\ell enjt}$	product $\times$ consumer-type utility of product j
ε_{ijt}	idiosyncratic consumer-preschool-year demand shock
mc_{jt}	preschools marginal costs
η_{jt}	preschools markup over marginal cost
π_{jt}	preschools variable profits
$\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3)$	Static parameters representing linear and non-linear utilities, and marginal costs
ξ_j	product level permanent demand shock
$\Delta\xi_{jt}$	product level transitory demand shock
ω_j	product level permanent marginal cost shock
$\Delta\omega_{jt}$	product level transitory marginal cost shock
$g(\boldsymbol{\theta})$	moments used in static estimation
$\mathcal{M}_{xjt}$	state summarizing public information available to preschool j at t in own-state x
$\mathcal{H}(\mathcal{M}_{xjt})$	set of simulated potential next states from $\mathcal{M}_{xjt}$
$\boldsymbol{\psi}$	dynamic parameters
$\epsilon(a)$	action contingent private payoff shock
$\mathbb{P}(a, \mathcal{M}_{xjt})$	strategy function assigning probability of action a in state $\mathcal{M}_{xjt}$
$\Pi(a; \mathcal{M}_{xjt})$	Preschool flow payoffs from action a in state $\mathcal{M}_{xjt}$
$\bar{\Pi}(\mathcal{M}_{xjt}; \mathbb{P})$	Preschool average flow payoffs in state $\mathcal{M}_{xjt}$ under strategy $\mathbb{P}$
$\boldsymbol{\Phi}$	basis of approximation for the value function
$\boldsymbol{\rho}$	weights on the approximation basis for the value function
κ	policy variable: schedule of quality add-on rates
$MCCA_{jt}$	policy variable: Maximum Child Care Allowance in center j county
Υ	set of policy variables

B ECE Policies in Pennsylvania

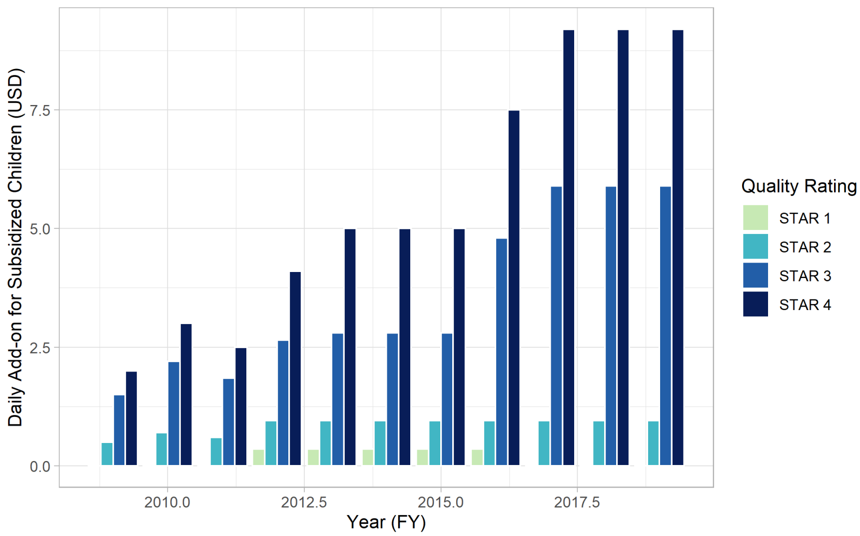
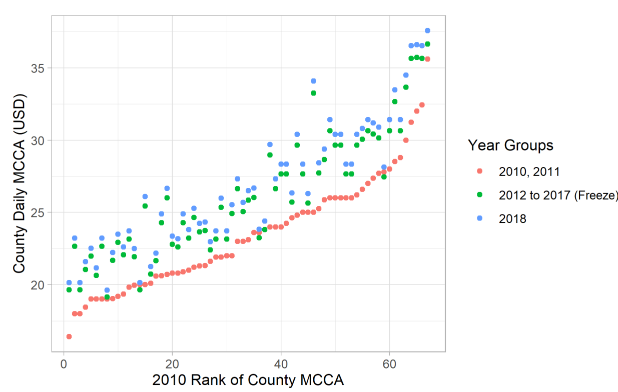


Figure B.1: Supply side: Add-on Rates by Quality Rating

Notes: This Figure presents the main supply side component of Keyston Stars. Programs at higher ratings receive a bonus per day, per subsidized child served, called add-on rate. The data comes from FOIA requests to the PA Department of Human Services.

(a) Copayment Schedule



(b) Maximum Child Care Allowance by County

Figure B.2: Components of the Child Care Works Demand Side Subsidy

Notes: The Figure describes the two components of the demand-side subsidies determining the price paid by sibsidized families. Copayment schedules are a function of Household Size and Income, and are presented on the left panel for a family of size 4. The source of the data are the Pennsylvania Bulletins recovered for the years 2010-2018. MCCA on the right panel vary by county, and are updated twice over the sample. The data comes from FOIA to the PA Department of Human Services.

C Centers Descriptive Facts - Supplementary Figures

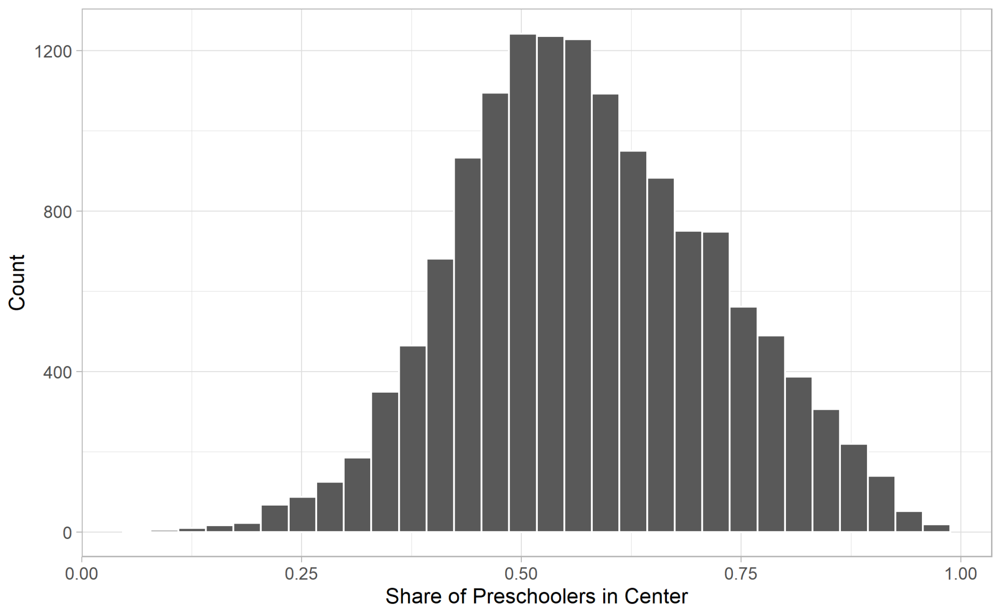
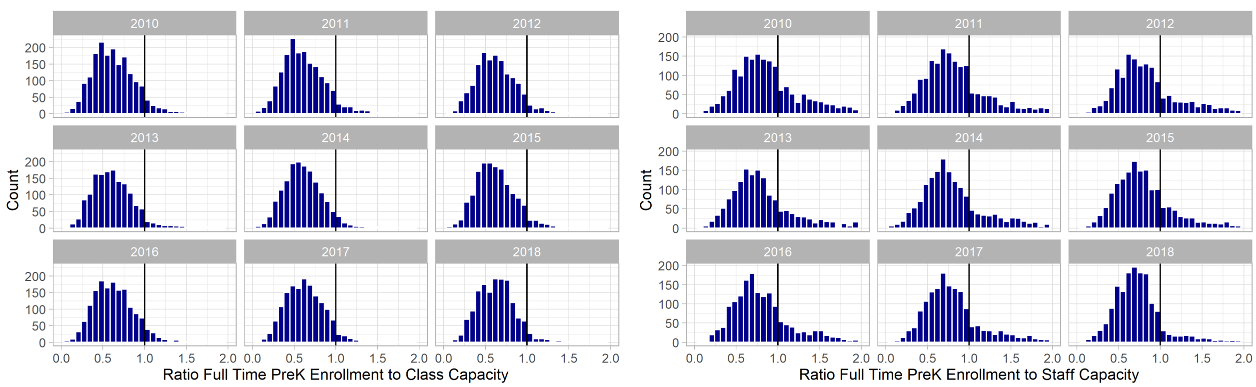


Figure C.1: Share of Preschoolers in Centers

Notes: Figure displays the distribution of the fraction of children enrolled belonging to the preschooler age group category (3-4 year olds). The remaining children are in majority toddler and infants, aged one to two.



(a) Class Size Maximum Requirements

(b) Staff Minimum Requirements

Figure C.2: Centers and Measures of Capacity Constraints

Notes: Left panel displays the distribution of the ratio of preschoolers enrollment to a capacity constraint based on the number of classrooms in the center, depicted by the black veritcal line. Pennsylvania mandates a maximum of 10 preschoolers per classroom. The right panel displays the ratio of preschoolers enrollment to a constraint based on staffing minima (2 adults per 10 children). Neither measures show a large share of centers with binding constraints.

D Stylized Facts - Supplementary Figures

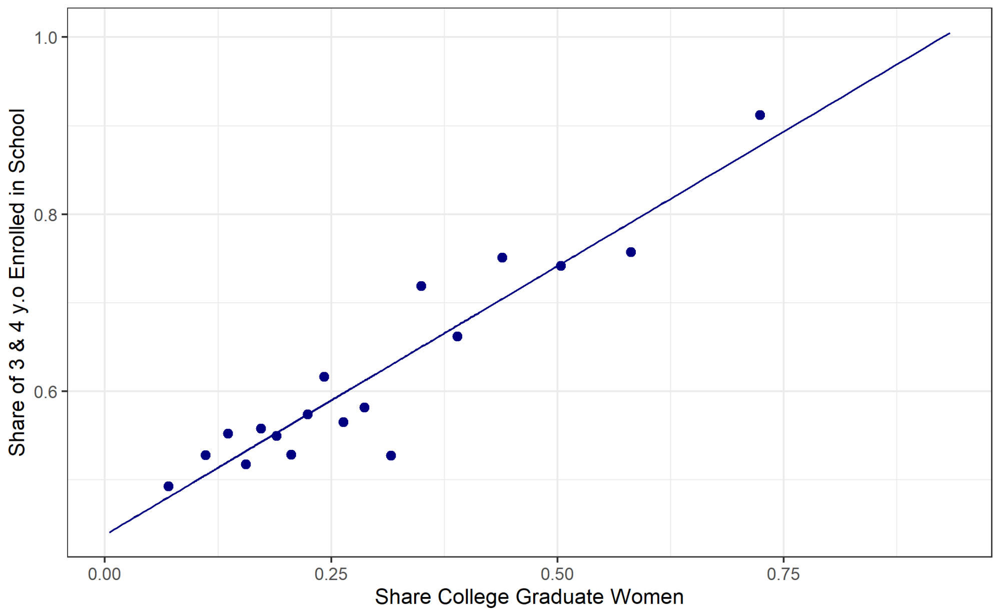
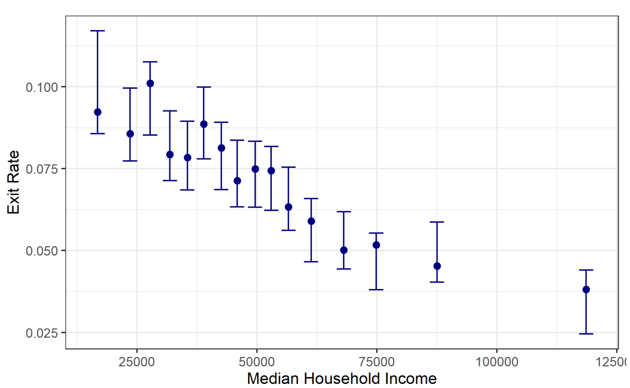
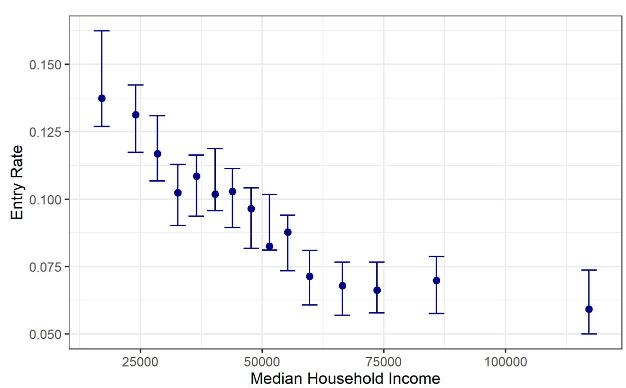


Figure D.1: Preschoolers Enrollment and College Graduate Women, Pennsylvania 2018

Notes: Figure displays preschool enrollment as reported in the American Community Survey 5 year estimates as a function the share of college graduate women in a census tracts. The number of bins is chosen using the (Cattaneo et al., 2019) data-driven optimal bin selection method.



(a) Exit Rate



(b) Entry Rate

Figure D.2: Centers Churn by Neighborhood Household Income

Notes: Left panel displays the average exit rates, right panel the average entry rates, over census tracts grouped in bins by median household income. The number of bins is chosen using the (Cattaneo et al., 2019) data-driven optimal bin selection method. 95% confidence intervals are displayed along with the bin average.

E Static Model - Supplementary Material

E.1 Simulated Families of Preschoolers

E.1.1 Census Demographics at Geographically Dis-aggregated Level

The estimation of the demand system proceeds by matching market shares from simulated families with observed market shares over preschools. These simulated families are constructed using ACS block-group data, in order to get the location contingent distribution of preschoolers, women educational attainment, and income distributions over time. ACS data provides geographically dis-aggregated information, but not all census variables are available at the block-group level. In particular, the breakdown of children' ages under five, and income distribution contigent on educational attainment, are not available at the block-group level. The breakdown of children' ages under five is available at the census tract level, and I assume that the share of 3 and 4 year olds among five year olds is uniform within the block groups constituting a given census tract. The smallest geography at which micro-data with both income and education is available is the Public Use Microdata Areas (PUMA). I assume that income is distributed according to a log-normal distribution. To recover the education contigent income distribution in a block group, I use information from PUMAs to form a prior, and the histogram of aggregate income in the block group as a signal to update this prior and get the posterior mean and variance of the lognormal distribution of income contingent on education.

The objects of interest are the parameters of the distribution of household income conditional on mothers' educational attainment in each block group b. I make the simplifying assumption that the variances of these conditional distributions are equal to the variances in the corresponding PUMA, and focus on computing the posterior on the means $(\mu_{b,H}, \mu_{b,L})$ of these conditional block group distributions. From the PUMAs, I get

$$\begin{bmatrix} \log y_H \\ \log y_L \end{bmatrix} = \begin{bmatrix} z_H \\ z_L \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_{p,H} \\ \mu_{p,L} \end{bmatrix}, \begin{bmatrix} \sigma^2_{p,H} & 0 \\ 0 & \sigma^2_{p,L} \end{bmatrix} \right)$$

Calling Σ^{μ} the covariance matrix of the distribution of $\begin{bmatrix} \mu_{p,H} \\ \mu_{p,L} \end{bmatrix}$ across PUMAs and years, I propose a bivariate normal prior for the parameters of interest:

$$\begin{bmatrix} \mu_{b,H} \\ \mu_{b,L} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_{p(b),H} \\ \mu_{p(b),L} \end{bmatrix}, \Sigma_\mu\right)$$

The signals consist in the aggregate distribution of log Household income z in block group b. Denoting by π_H the share of mothers with a college degree in block group b, I can express z as a mixture of the log income distributions of the two education groups of households and follows:

$$z \sim \underbrace{(\pi_H, 1-\pi_H)}_{\pi'}.\mathcal{N}\left(\begin{bmatrix} \mu_{p,H} \\ \mu_{p,L} \end{bmatrix}, \begin{bmatrix} \sigma^2_{p,H} & 0 \\ 0 & \sigma^2_{p,L} \end{bmatrix}\right) \sim \mathcal{N}(\pi'\mu_b, \underbrace{(\pi_h\sigma_{p,H})^2 + ((1-\pi_H)\sigma_{p,L})^2}_{\sigma^2_p})$$

I can know write the posterior distribution on μ_b given the signal z and our prior

$$\begin{aligned} p(\mu_b|z,\theta) \propto \ell(z|\theta)p(\theta) &\propto \exp\left(-\frac{\sum_i^N (z_i - \pi'\mu_b)^2}{2\sigma^2_p}\right)\exp\left(\frac{-1}{2}(\mu_b-\mu_p)'\Sigma^{-1}_\mu(\mu_b-\mu_p)\right) \\ &\propto \exp\left(\frac{N\bar{z}\pi'\mu_b}{\sigma^2_p} - \frac{N}{2\sigma^2_p}\mu_b'\pi\pi'\mu_b - \frac{1}{2}[\mu_b'\Sigma^{-1}_\mu\mu_b - 2\mu_p'\Sigma^{-1}_\mu\mu_b]\right) \\ &\propto \exp\left(\left[\frac{N\bar{z}}{\sigma^2_p}\pi' + \mu_p'\Sigma^{-1}_\mu\right]\mu_b - \frac{1}{2}\mu_b'\left[\frac{N}{\sigma^2_p}\pi\pi' + \Sigma^{-1}_\mu\right]\mu_b\right) \end{aligned}$$

Which we can term by term identify with a bivariate normal distribution of the form

$$\exp\left(\frac{-1}{2}(\mu_b-\mu_\theta)'\Sigma^{-1}_\theta(\mu_b-\mu_\theta)\right)$$

Where

$$\begin{aligned} \Sigma^{-1}_\theta &= \frac{N}{\sigma^2_p}\pi\pi' + \Sigma^{-1}_\mu \\ \mu_\theta &= \left(\left[\frac{N\bar{z}}{\sigma^2_p}\pi' + \mu_p'\Sigma^{-1}_\mu\right]\left[\frac{N}{\sigma^2_p}\pi\pi' + \Sigma^{-1}_\mu\right]^- 1\right)' \end{aligned}$$

The two components of μ_θ give us the posterior means of the two conditional distributions of log household income in the block group b. Given the parameters of these distributions, for each block-group b I know have a measure of the number of preschoolers, the share of college educated women, and household income distributions for households with college and non college educated women.

E.2 Numerical Integration over Simulated Families

I use quadrature to construct weights associated with the log-normal distributions of income derived in the previous paragraph. While estimation of differentiated products demand models can use either Monte Carlo or quadrature rules to generate agent nodes and weights, in the case of spatially differentiated product the locations already introduce a high number of spatial nodes (as many as block groups). Quadrature rules keep the number of simulated nodes manageable by only requiring of few number of well chosen income values to take a weighted sum over. Quadrature rules require the integrand to be a continuous function. The subsidies on the demand side introduce a discontinuity at the income eligibility threshold for the market share function. Consequently, I split the numerical integration in two intervals and approximated separately the integral over $y_i \in [0, 2 * FPL[$ and $y_i \in [2 * FPL, \infty[$ where two times the Federal Poverty Line corresponds to the thresholds for subsidies in the data. In counterfactuals involving a different income eligibility threshold, new income nodes and weights are generated.

I find that Gauss-Legendre quadrature provides a good approximation of the integral of market shares functions over a truncated normal distributions on semi-bounded intervals. Consequently, I use Gauss-Legendre quadrature which for the 3 types of intervals, [0,b], [a,∞], and [a,b] yields: For semi-closed intervals $[0, \bar{y}]$:

$$\int_0^{\bar{y}} s(y) \frac{1}{y\sigma\sqrt{2\pi}} \exp(\frac{-1}{2\sigma^2}(\log y - \mu)^2) dy$$

$$y_n = \frac{\bar{y}}{2}(v_n^{Legendre} + 1)$$

$$w_n = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{v_n^{Legendre} + 1} \exp\left(\frac{-1}{2\sigma^2}\left(\log\left(\frac{\bar{y}(v_n^{Legendre} + 1)}{2}\right) - \mu\right)^2\right) w_n^{Legendre}$$

For the closed intervals [a,b]:

$$\int_a^b s(y) \frac{1}{y\sigma\sqrt{2\pi}} \exp(\frac{-1}{2\sigma^2}(\log y - \mu)^2) dy$$

$$y_n = \exp\left(\frac{1}{2} \log \frac{b}{a}(v^{Legendre} + 1) + \log a\right)$$

$$w_n = \frac{1}{2} \log \frac{b}{a} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-1}{2\sigma^2}\left(\frac{1}{2} \log \frac{b}{a}(v^{Legendre} + 1) + \log a - \mu\right)\right) w_n^{Legendre}$$

Lastly, for semi-closed intervals $[a, \infty]$:

$$\int_a^{\infty} s(y) \frac{1}{y\sigma\sqrt{2\pi}} \exp(\frac{-1}{2\sigma^2}(\log y - \mu)^2) dy$$

$$y_n = \frac{2a}{v_n^{Legendre} + 1}$$

$$w_n = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{v_n^{Legendre} + 1} \exp\left(\frac{-1}{2\sigma^2}\left(\log\left(\frac{2a}{v^{Legendre} + 1}\right) - \mu\right)^2\right) w_n^{Legendre}$$

E.3 Markets

Market	Preschoolers	Subsidized	Non Subsidized	Mean Household Income	Mean Daily Price	STAR 1	STAR 2	STAR 3	STAR 4	Block Groups
Abington Hatboro Horsham	8626	3044	5582	121950	49	39	6	7	18	262
Allentown Bethlehem Easton	14919	7062	7857	87313	34	95	42	28	35	444
Altoona	2355	1357	997	64555	29	11	8	1	15	101
Cannonsburg Peters	2254	705	1549	129025	45	12	1	2	7	61
Carlisle	1389	602	787	91221	34	6	0	1	7	37
Central Bucks North Penn	7803	2300	5504	132405	46	45	9	8	26	202
Chambersburg Shippensburg	2470	1165	1306	86939	29	14	4	1	3	58
Chester	8960	3703	5257	113769	45	43	15	5	16	306
Erie	5833	3407	2426	69728	35	14	15	12	28	217
Freeport Burrell Newkensington	3037	1438	1598	91589	34	16	10	4	2	121
Greensbrug	2538	1226	1312	84508	36	7	6	2	9	98
Hanover	2180	955	1225	85116	32	10	3	1	3	61
Harrisburg	11883	5651	6232	90532	39	77	39	18	26	325
Johnstown	1960	1212	748	61925	27	11	4	4	3	102
Lancaster	9888	4395	5493	88650	39	31	9	9	34	245
Laurel Highlands	2311	1438	873	62578	30	10	7	3	4	86
Levittown	9497	3439	6058	118470	45	44	20	16	21	290
Mars Butler Seneca	3743	1410	2333	116977	38	23	4	3	6	99
Moon Sewickley	1739	613	1126	114464	44	6	1	2	3	55
Newcastle	1985	1136	849	66317	35	3	3	5	4	78

Market	Preschoolers	Subsidized	Non Subsidized	Mean Household Income	Mean Daily Price	STAR 1	STAR 2	STAR 3	STAR 4	Block Groups
Norristown	7712	2478	5234	138114	48	48	11	8	15	226
Penn Trafford Norwin	2851	1196	1654	101744	35	16	2	2	5	92
Philly Center	17912	12595	5317	53792	38	112	26	23	14	500
Philly Northeast	18300	10755	7544	70585	38	115	34	19	24	448
Philly Northwest	16527	9665	6862	77251	38	144	37	19	21	520
Philly South	9447	5618	3829	82145	44	82	11	16	18	387
Philly Southwest	22681	11763	10917	96522	40	240	59	31	44	808
Pitt East	10797	6030	4767	79441	38	84	23	13	38	562
Pitt North	8476	3554	4922	112585	42	40	10	5	16	321
Pitt South	11607	5884	5724	86686	41	68	12	4	18	545
Pottstown Phoenixville	7446	2469	4977	125920	44	29	15	9	31	179
Reading	7239	3916	3323	76616	37	26	18	6	15	187
Rochester Aliquippa	4325	2039	2286	89320	33	24	11	8	5	160
Scranton	4590	2482	2108	75681	33	18	21	8	15	181
State College	1896	883	1013	108554	41	13	6	6	9	82
Stroudsbourg Pocono	3374	1553	1821	96905	32	44	0	4	7	96
Westchester Greatvalley Dorringtown	8924	2584	6340	146192	49	50	18	6	25	211
Wilkes Barre	5860	3199	2661	71579	30	33	17	13	9	230
York	6432	3164	3268	86663	33	15	18	7	13	201

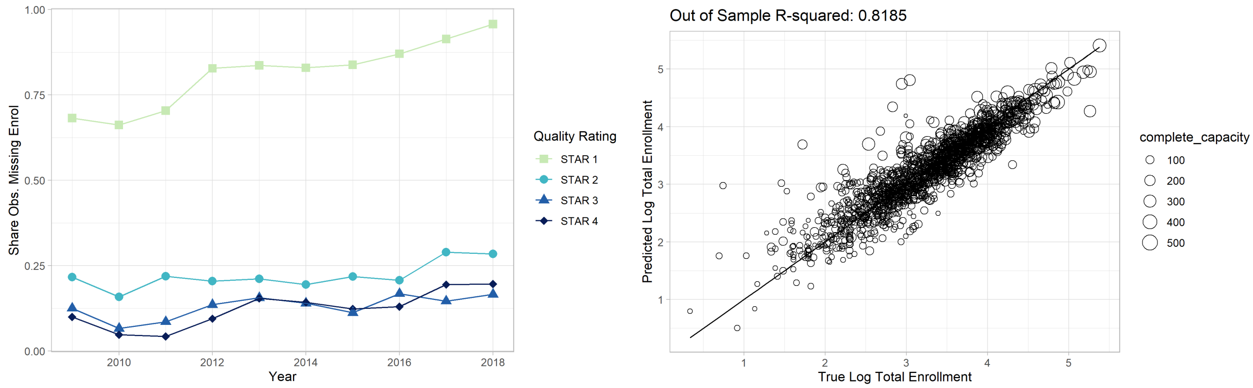
E.4 Demand Estimation with Missing Market Shares

General Assumptions. Enrollment data is provided as Full-Time equivalent level, so the model abstracts away from the full-day/part-day margin and assumes that all parents send their child to preschool full day. The model also abstracts away from sibling effects and household with multiple children for now.

Missingness Patterns in Enrollment Data. Submitting annual enrollment reports is required for centers who want to claim eligibility for financial support offered as part the Pennsylvania's QRIS. However, STAR 1 centers are not eligible to any support in this system, and therefore have little reasons to submit reports. As a result, the fraction of STAR 1 centers reporting enrollment is low in the data, while it is very high for STAR 2-4 centers, as can be seen on Figure E.1a. The STAR 1 centers that do report enrollment tend to upgrade to higher quality ratings later. Missing enrollment information is a challenge for differentiated product demand estimation, as markets shares of all of the products are required to inverse the demand system and recover the product specific utility components $\boldsymbol{\delta}$. While a common approach in empirical papers is to simply drop products with missing quantities, I am reluctant to do so as there are many such centers, implying that dropping them would likely misrepresent the degree of spatial competition in the static model.[49] In addition, dropping these centers would result in the loss informative data points for dynamic estimation. Instead, I propose to split the data into two groups of products: the first one, denoted by $\mathcal{J}^{obs}$, contains all providers that ever report their enrollment. For these centers, the absence of reporting is rare and missing enrollment information can be credibly imputed using machine learning, as shown in Figure E.1b. The second group of products, denoted by $\mathcal{J}^{miss}$, contains centers which never report their enrollment. Individual imputations for these centers would be a stretch, but it is possible to get a good sense of how many children in aggregate go to these centers in each market. In estimation, I therefore treat this products as an aggregate group for which a unique market share is used. I present this approach below.

Assumption for Missing STAR 1. For preschools $j \in \mathcal{J}^{miss}_{mt}$, the lack of product-level shares precludes the identification of a product-level ξ_{jt}. Instead **we make the assumption** $\xi_{jt} = \xi_{\o mt} \forall j \in \mathcal{J}^{miss}_{mt}$. In addition, as these centers are never rated higher than STAR 1, the $\beta^(sr)$ terms drop from the utility specification, which becomes

[49] Econometrics papers on demand estimation with zeroes in market shares such as Dubé et al. (2021) and Gandhi et al. (2020) contain a discussion on the drawbacks of dropping products in demand estimation.



(a) Share of Missing Enrollment

(b) XGBoost Out of Sample Prediction, $\mathcal{J}^{obs}$

Figure E.1: Missing Enrollment Diagnostic, and Prediction for $\mathcal{J}^{obs}$

*Notes:*The left panel show the fraction of centers in each given year not reporting their enrollment, broken doan by quality rating. THe right panel shows the out of sample performance of an XGBoost regression model trained to predict enrollment for the preschools in $\mathcal{J}^{obs}$, based on lead, lag center characteristics, and on nearby demographics and competitors. The prediction R^2 is displayed above the figure.

$$
\begin{aligned}
u_{i\ell ejt} = & \underbrace{\xi_{\o mt}}_{\delta_{\o mt}} \\
& + \underbrace{\beta_H^{(0)} \mathbf{1}\{e_i = H\} + \left(\alpha_1^e + \frac{\alpha_2^e}{y_i}\right) p_{ijt} + \lambda^e d_{\ell j} + \varepsilon_{ijt}}_{\mu_{ie\ell jt}}
\end{aligned} \tag{16}
$$

The benefits of this approach relative to dropping the school is made apparent by this specification. While the market share is missing, the data still contains informative variables such as the price and location of the center. This information is exploited by the model in the household-product specific term $\mu_{ie\ell jt}$. However, the absence of center specific enrollement makes it impossible to identify a product-specific ξ, which results in assuming a common $\xi_{\o mt}$ for all products in $\mathcal{J}_{mt}^{miss}$. Following Elliott et al. (2021), utilities of products in $\mathcal{J}_{mt}^{miss}$ can be expressed as a function of mean characteristics of the group and product characteristics using

$$
\begin{aligned}
\delta_{\o mt} &= \xi_{\o mt} \\
\mu_{e\ell n\o mt}(\beta_{cap}, \boldsymbol{\theta}_2) &= \log \sum_{j \in \mathcal{J}_{mt}^{miss}} \exp \mu_{e\ell njt}(\boldsymbol{\theta}_2)
\end{aligned}
$$

Then we can write the total share of the group $\mathcal{J}_{mt}^{miss}$ demanded by households of type (e, ℓ, y_n) as

$$s_{\ell n \o t}^{e}(\boldsymbol{\delta}_t, \boldsymbol{\theta_2}) = \sum_{j \in \mathcal{J}_{mt}^{miss}} \frac{\exp(\delta_{\o mt} + \mu_{e\ell njt}(\boldsymbol{\theta}_2))}{1 + \sum_{j \in \mathcal{J}_{mt}^{obs}} \exp(\delta_{jt} + \mu_{e\ell njt}(\boldsymbol{\theta}_2)) + \sum_{j \in \mathcal{J}_{mt}^{miss}} \exp(\delta_{\o mt} + \mu_{e\ell njt}(\boldsymbol{\theta}_2))}$$
$$= \frac{\exp(\delta_{\o mt} + \mu_{e\ell n\o mt}(\boldsymbol{\theta}_2))}{1 + \sum_{j \in \mathcal{J}_{mt}^{obs}} \exp(\delta_{jt} + \mu_{e\ell njt}(\boldsymbol{\theta}_2)) + \exp(\delta_{\o mt} + \mu_{e\ell n\o mt}(\boldsymbol{\theta}_2))}$$

Given this expression, we are back to the BLP set-up where the contraction mapping is a contraction. Given a vector $\left\{(s_{jt})_{\mathcal{J}_{mt}^{obs}}, s_{\o mt}\right\}$ we can recover a unique vector $\left\{(\delta_{jt})_{\mathcal{J}_{mt}^{obs}}, \delta_{\o mt}\right\}$ for any guess of parameters $\boldsymbol{\theta}_2$.

E.5 Choice of Instruments

Estimation of the non-linear parameters relies on the following exclusion restriction:

$$\mathbb{E}[\Delta\xi_{jt} z_{jt}] = 0$$

Instruments are constructed following the differentiation instruments method introduced by Gandhi and Houde (2019) and adapted to spatially differentiated setting. Consider a characteristic of the preschool x_{jt} which is exogenous within a market period, i.e is given when the shock $\Delta\xi_{jt}$ is revealed to the center. These characteristics includes the presence of the center in the market, its current STAR rating and its capacity. Isolation of the center within it's local market can be measured as

$$z_{jt}^{(x)} = \sum \frac{(x_{jt} - x_{kt})^2}{d_{jk}}$$

I first construct instruments $\{z_{jt}^{(x)}, z_{jt}^{(x)} \bigotimes z_{jt}^{(x')}, z_{jt}^{(x)} \bigotimes dem_{jt}\}$ where the demographics dem_{jt} include statistics on household structure, female labor supply, income and racial mix in the neighborhood around j at t. Then, following Backus et al. (2021), I use these instruments to predict a measure of exogenous price $\hat{p}_{jt} = \mathbb{E}[p_{jt}|x_{jt}, z_{jt}]$. Intuitively, this exogenous price can be understood as capturing as much of the first-stage variation as possible in a single instrument. Following the same approach, it is used to construct the additional instruments $\{z_{jt}^{(\hat{p})}, z_{jt}^{(\hat{p})} \bigotimes z_{jt}^{(x)}, z_{jt}^{(\hat{p})} \bigotimes dem_{jt}\}$. This procedure generates more than a hundred instruments. I then select the 20 instruments which are the most significant in the first-stage regression of price on exogenous characteristics and excluded instruments. The regressions of price and market share on these excluded instruments, centers characteristics, and center fixed effects yields a F-statistic for excluded instruments around 30.

E.6 Additional Moments

In addition to the moments described in the main text, estimation uses additional micromoments constructed from survey data to inform the parameters governing the heterogeneous aspects of parents' for preschools' characteristics. The NSECE surveys parents of young children, and in particular records the child care costs for parents using specific arrangements and the distance traveled to these arrangements, which I convert in travel time. In addition, the NSECE contains household members demographics. This survey allow me to form the expectation of a preschool characteristic y such as price or distance conditional on household type and on using preschool. The corresponding model conditional expectation is given by:

$$v_e^y(\boldsymbol{\theta}_2) = \sum_{i \in e} \frac{w_i}{\sum_{i \in e} w_i} \sum_{\mathcal{J}} \frac{s_{ij}(\boldsymbol{\theta}_2) y_{ij}}{1 - s_{i0}(\boldsymbol{\theta}_2)}$$

Denoting by $v_e^{y,survey}$ the survey counterpart, I define an additional moment which is stacked to the GMM objective:

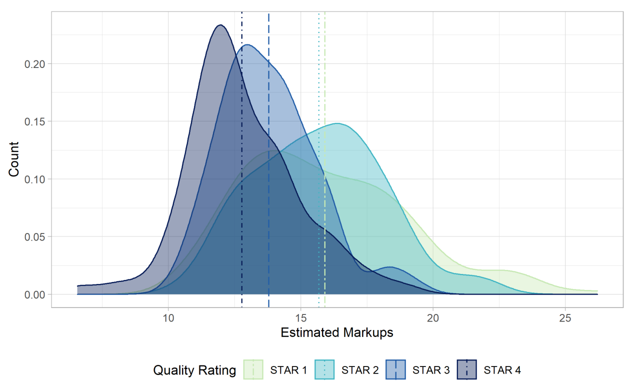
$$g_e^y(\boldsymbol{\theta}_2) = \frac{v_e^y(\boldsymbol{\theta}_2) - v_e^{y,survey}}{v_e^{y,survey}}$$

The analytic gradient correpsonding to the additional moment conditions is similarly stacked to the gradient of the GMM objective. As is standard in the estimation of these models using different sources of moments, the co-variance between the IV-GMM, the private-pay shares moments, and the additional micro-moments is assumed to be zero, leading to a co-variance matrix which is block diagonal.

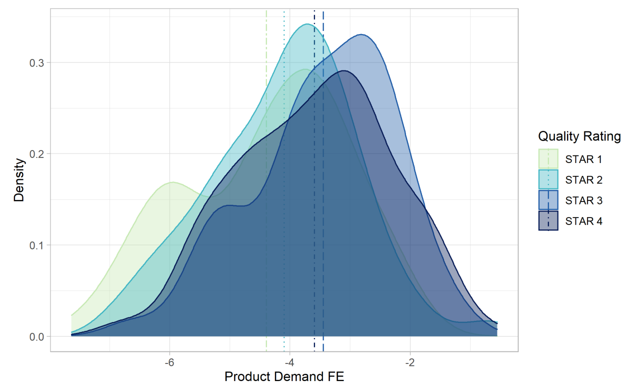
E.7 Model Fit - Additional Figures

E.8 Solution Method for Static Equilibrium

The estimated static model can be used to compute the equilibrium corresponding to a new market state $\mathcal{M}_t$. The equilibrium outcomes of interest are prices $\{\boldsymbol{p}\}$, market shares $\{\boldsymbol{s}\}$, and the resulting profits $\{\boldsymbol{\pi}\}$. Estimation and counterfactuals require computing a large number of these updated equilibria, following changes in $\mathcal{M}_t$ that involve: alteration in the demographic composition of a market, in the market structure, and in the demand and cost shocks (to construct the grid of N points to approximate the value function, and to construct the set of possible next states following these N points). It also involves for the counterfactuals changing the add-on rates or copayment schedules



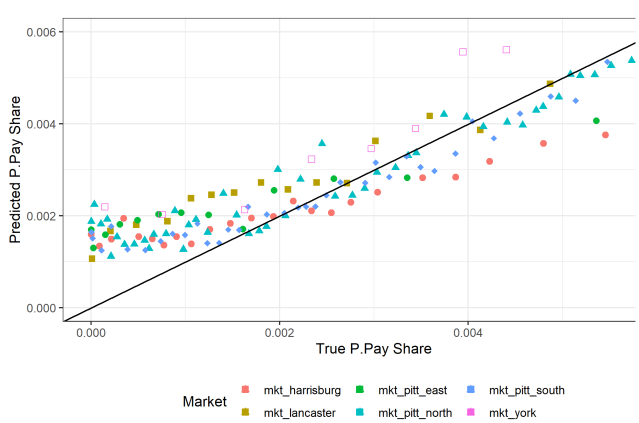
(a) Markups in 2015



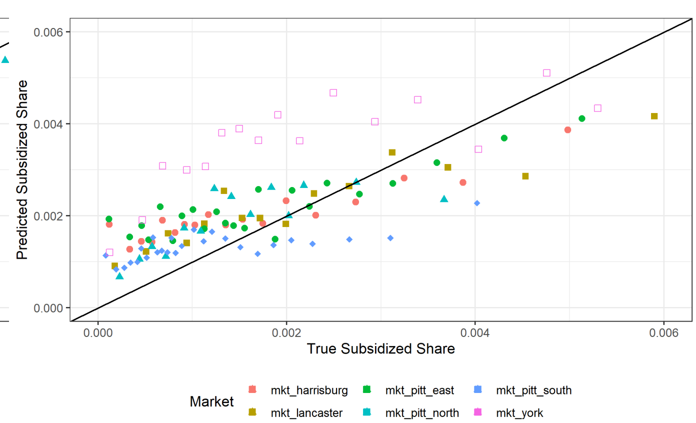
(b) Demand Unobservable Fixed Effect by Rating in 2018

Figure E.2: Markups and Unobserved Quality by STAR Rating

Notes: Additional figures for static demand and supply estimation. The left panel displays the distribution of markups in 2015 by STAR rating. The counter-intuitive fact that higher ratings charge lower markups can be explained by the large add-on rates they benefit from, which discourage charging higher markups. The right panel shows the preschool fixed effect on unobserved demand shock by 2018 STAR rating. This graph suggest selection forces at play, where unobservably better centers also tend to upgrade to higher ratings.



(a) Private Pay Market Shares, Data and Model



(b) Subsidized Market Shares, Data and Model

Figure E.3: Demand Model Fit Across Markets

Notes: Binscatter plot of the private pay (left panel) and subsidized (right) market shares predicted by the model compared to those reported by centers in the data. .

which together lead to a new policy environment Υ. To perform efficiently these computations, I first follow the recommendation of Conlon and Gortmaker (2020) and reformulate the equilibrium as a fixed point in terms of modified markups which I denote by $\{\boldsymbol{\zeta}(\boldsymbol{p})\}$ instead of $\{\boldsymbol{\eta}(\boldsymbol{p})\}$ for the BLP markups. These two functions of the price vectors differ in general, but coincide at equilibrium prices. Morrow and Skerlos (2011) find that solving for equilibrium prices using the Picard iteration routine $\boldsymbol{p}^{t+1} \leftarrow \boldsymbol{mc} + \boldsymbol{\zeta}(\boldsymbol{p}^t)$ is fast and reliable, while $\boldsymbol{p}^{t+1} \leftarrow \boldsymbol{mc} + \boldsymbol{\eta}(\boldsymbol{p}^t)$ may not converge. To further accelerate the solution routine for static equilibrium prices, I replace the Picard iteration scheme with Anderson acceleration (Anderson, 1965) which I find greatly decreases the number of iterations needed to reach convergence. The presence of the add-on rates policy κ slightly modifies the expression for $\boldsymbol{\zeta}$, with specific contributions from subsidized and non subsidized parents. For $k \in \{sub, nosub\}$ denote the two terms of $\frac{\partial s_j}{\partial p_j}$ as:

$$\sum_{i \in k} w_i \alpha_i \frac{\partial s_{ij}}{\partial p_j} s_{ij} = \Lambda_k \text{ and } \sum_{i \in k} w_i \alpha_i \frac{\partial s_{ij}}{\partial p_j} s_{ij}^2 = \Gamma_k \text{ such that } \frac{\partial s_{kj}}{\partial p_j} = \Lambda_k - \Gamma_k$$

Then the function $\zeta(;\kappa)$ is defined as

$$\boldsymbol{\zeta}(\boldsymbol{p};\kappa) = \Lambda^{-1}\Gamma(\boldsymbol{p} - \boldsymbol{mc}) - \Lambda^{-1}\boldsymbol{s}(\boldsymbol{p}) - w_{sub}\Lambda^{-1}(\Lambda_{sub} - \Gamma_{sub})\boldsymbol{\kappa}(\boldsymbol{sr})$$

where Λ, Γ are the contributions of the derivative of market shares aggregated over subsidized and non-subsidized parents, and $\boldsymbol{sr}$ is the vector of STAR ratings in the market determining the add-on rates each school receives for its subsidized enrollment.

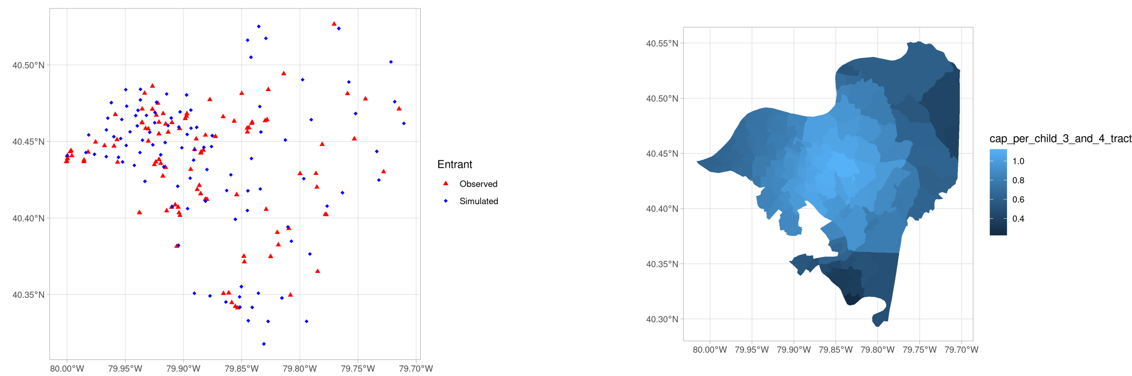
F Dynamic Model - Supplementary Material

F.1 Potential Entrants Simulation

The dynamic game in this paper involves two types of players, incumbents and potential entrants. Potential entrants are short-lived: if they decide to exit, they are taken out of the game altogether, while if they enter they become long-lived incumbents. Entries are observed in the data, but there might in principle be potential preschool operators who have decided not to join the market and have not generated any observed data point. Setting-up the game thus requires taking a stance on the number and characteristics of unobserved potential entrants, who have not entered in the data but could have entered under different market conditions. The presence of this modeling choice is standard in the literature on entry games, and papers tend to rely on a number of potential entrant in a market which is an increasing number of observed entries (Singleton, 2019; Caoui et al., 2022). The choice a specific number of potential entrants matters for estimation. As the number of observed entries remains the same, higher number of potential entrants will mechanically lead to larger estimated entry costs (Caoui et al., 2022). But the ultimate purpose of the model is to run counterfactuals. In a counterfactual, a larger number of entrants implies a greater potential for additional entries. These two effects counterbalance each other, and the ultimate impact of the choice of the number of potential entrants on the results of the policy simulations from the model is a priori unclear.

I fix the number of potential entrants in a market $\times$ year to twice the maximum number of entries observed in that market over the panel period. The fixed characteristics of potential entrants $\{\xi_j, \omega_j\}$ are drawn from the distribution of these unobservables over centers operating in the market. This modeling choice is made for simplicity, and is ignoring the selection effect due to entrants' knowledge of these unosbervables when making their entry decisions. Future robustness exercises could include drawing these characteristics from a a truncated distribution.

The model assumes that potential entrants are tied to a specific location in the market. This modeling choice reduces the size of entrants' action set by removing the location choice aspect from their entry decision. I simulate potential entrants on the map based on census tracts population of preschoolers and density, using a Poisson model trained on observed entries. When drawing a location for a potential entrant, I use weights generate from this model but oversample low-access census tracts to open up the possibility for an expansion of the supply in these areas during counterfactuals. The outcome of this spatial simulation process is shown on Figure F.1 below:



(a) Observed and Simulated Entrants, All Years

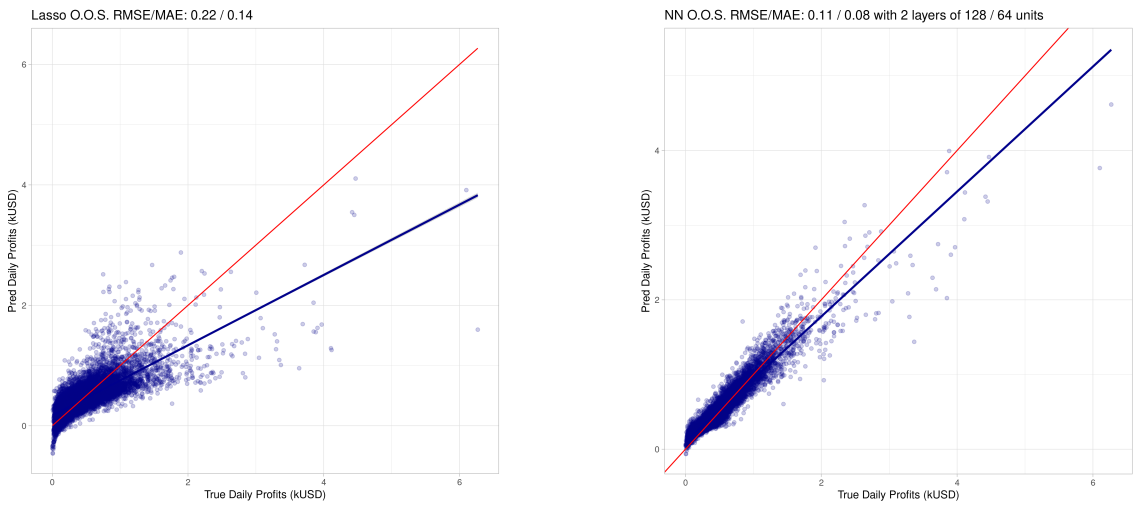
(b) Capacity Per Child

Figure F.1: Observed and Simulated Potential Entrants in East Pittsburgh

Notes: The Figure shows the result from the spatial simulation procedure for potential entrants described above. Potential entrnats in blue are assigned to similar locations as observed entrants in red, as these locations are drwan based on preschooler population, but low-access locations, such as the southern tip of the map, are over-sampled.

F.2 Basis $\mathbf{\Phi}$ for Value Function Approximation

The estimation of the dynamic parameters relies on a parametric approximation of the ex-ante value function to bypass the curse of dimensionality. The value function can also be written as the discounted sum of flow profits along the optimal policy path, motivating a choice of basis as a collection of functions of the state that provide a good approximation to variable profits (Sweeting, 2013; Beresteanu and Ellickson, 2019). I proceed in 2 steps to construct $\mathbf{\Phi}$. The first stage consists in translating the state $\mathcal{M}_{xjt}$ in a wide set of variables $\boldsymbol{X}(\mathcal{M}_{xjt})$ describing the demographics and market structure as richly as possible. These variables include distance weighted averages of demographics, observable and unobservable competitors characteristics, and counts of these variables in distance bands around j. The second stage aims at constructing non linear functions of the of this initial representation of the state that may be more even informative on the value function. To do so, I rely on a neural network taking $\boldsymbol{X}(\mathcal{M}_{xjt})$ as input and trained to predict $\pi(\mathcal{M}_{xjt}; \boldsymbol{\theta}_2, \Upsilon_t)$. The neural network can be trained using as much data as desired, because it is always possible to generate new perturbed states $\tilde{\mathcal{M}_{xjt}}$ and to compute the resulting variable profits in these new states with the static equilibrium routine described in Appendix E.8. In other words, this step consists in teaching the static model to a neural network by simulating many static equilibria. The improvement in terms of fit of the variables profits from using a neural network instead of simply relying on the initial bases $\boldsymbol{X}(\mathcal{M}_{xjt})$ is shown on Figure F.2. The neural network vastly outperforms the linear model, halving the root mean square error. The benefits form this extra stage are twofold. First, the last hidden layer of the network provides a low-dimensional, non linear representation of the state $\mathcal{M}_{xjt}$ and can be used to improve the basis $\boldsymbol{\Phi}$. Second, the network as a whole provides a much faster way to compute variable profits in many simulated states than the static equilibrium routine, at the cost of a small approximation error.



(a) Lasso

(b) Neural Network, 2 Layers

Figure F.2: Approximation of Variable Profits, Out of Sample Fit

Notes: Models are trained on 80% of the data and 20% is set aside to test model performance. These two graphs represent model fit on this set-aside data. The data used to train and test the models is generated by taking the observed profits, and adding perturbed states and the resulting equilibria using the algorithm presented in Appendix E.6. .

Figure F.3 summarizes the value function approximation and the role of the neural network in bypassing the large state space problem when solving the dynamic game.

F.3 Likelihood Function

The variant of the NPL estimator used to recover the dynamic parameters relies on minimizing a maximizing a pseudo-likelihood function at each iteration, for a given guess of parameters and policy functions. Out of the grid of N points used to compute the value function approximation, denote by N^{Data} the set of points (and by extension the number of these points) corresponding to observed centers in the data. The other points are perturbed states used to improve the out-of-sample fit of the value function approximation. Denote by a_n^{Data} the action taken in state $n \in N^{Data}$. The negative log-likelihood which is to be minimized at iteration i characterized by strategies $\times$

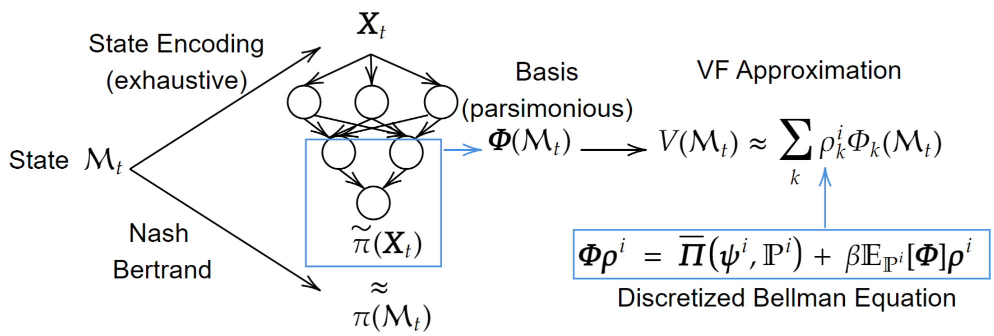


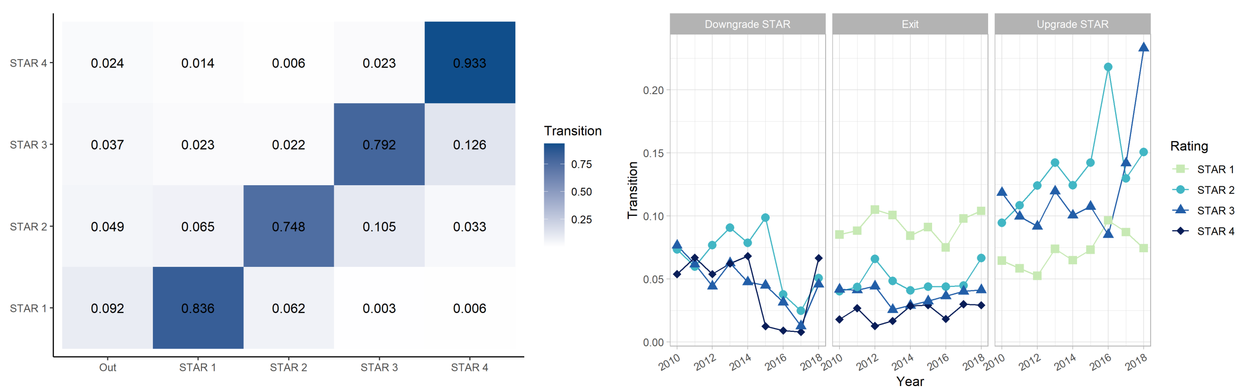
Figure F.3: Objects used in Approximation to Solve the Game

Notes: Summary of the value function approximation embedded in the parametric policy iteration used to solve the game. $\bar{\mathbf{\Pi}}(\boldsymbol{\psi}, \mathbb{P})$ denotes the expected flow profits given the optimal strategy and parameters at iteration i.

parameters guess $(\mathbb{P}^i, \boldsymbol{\psi}^i)$ can be expressed as:

$$\ell(\boldsymbol{\psi}; \mathbb{P}^i, \boldsymbol{\psi}^i) = -\frac{1}{N^{Data}} \sum_{n \in N^{Data}} \log(\mathbb{P}(a_n^{Data} | n, \mathbb{P}^i, \boldsymbol{\psi}^i))$$

F.4 Additional Figures - Model Fit



(a) Frequency of Action by State

(b) Action Probabilities over Time

Figure F.4: Dynamic Decisions

Notes: Left panel: Average of action taken by preschool STAR rating over years 2010-2018. Right panel: average of most frequent moves, downgrade one rating, exit, and upgrade one rating, over years.

Table 9: Dynamic Model Fit

Rating	Action	Proba. Data	Proba. Model
		Upgrade One Rating	
Star 1	Star 2	0.06	0.06
Star 2	Star 3	0.09	0.09
Star 3	Star 4	0.11	0.11
		Exit	
Star 1	out	0.10	0.11
Star 2	out	0.06	0.02
Star 3	out	0.05	0.01
Star 4	out	0.02	0.00

Notes: Results from the Nested Pseudo Likelihood estimator. Column Proba. Data shows the empirical probability of taking exiting or upgrading one rating in the data used in estimation. Column Proba. Model shows the averages for exit and upgrade one rating pair of the equilibrium policy function.

G Counterfactuals - Supplementary Material

G.1 Initial Probabilities in Counterfactual Policy Environments

The PPI approach used to solve the game requires a starting point for preschools' strategies. In estimation, I use the conditional choice probabilities, but in counterfactual environments these probabilities are unlikely to provide a great starting point, as new policies can shift the value of operating in different neighborhoods and at different quality ratings. I initialize the strategies using a similar approach to Sweeting (2013). Given estimates of the transition and fixed costs, I estimate a simple multinomial logit model where I assume that school's next states are only functions of the payoffs (variable profits minus estimated transition and fixed costs) that they would get in each state. This amounts to treating preschools as both nonstrategic and myopic agents who do not take into account neither their competitors strategies nor continuation values when evaluating potential actions. This multinomial logit yields a positive coefficient on this payoff variable.[50] A counterfactual environment Υ can modify these initial probabilities by either changing variable profits in each state (for instance through a change in demand subsidies or add-on rates) or change the costs associated with transitions or operating a center.

[50] I treat entry separately and in this version of the paper use the same probabilities of entry as in the data as a starting point.

G.2 Importance Sampling in Spatial Games with Many Players

When solving the dynamic game, expectations at iteration i are computed by weighted importance sampling. Weights of each potential next state in the grid are updated according to the new iterate of the policy function $\mathbb{P}^i$. When solving for counterfactuals in particular, both the contribution of player j's strategy and the contribution of the strategies of player j's competitors are updated every iteration. The contribution of competitors $\Pi_{k \neq j} \mathbb{P}_k^i$ can create a stability issue, as small changes in each of the individual players' probabilities can lead to large changes in the weight of a state whose realization depends on the actions of all of the competitors. To mitigate this concern, I use the spatial nature of the static competition framework and assume that player j only takes into account the effect of competitors in a radius of consideration $\mathcal{R}_{jt} \in \mathcal{J}_{mt}$. This assumption follows from the fact that given the importance of distance in shaping parents' demand for early education, distant competitors' actions are unlikely to have any impact on payoffs. In the counterfactuals, I take $\mathcal{R}_{jt}$ to be the union of the 5 closest competitors and the set of competitors in a $1000m$ radius.

G.3 Additional Figures

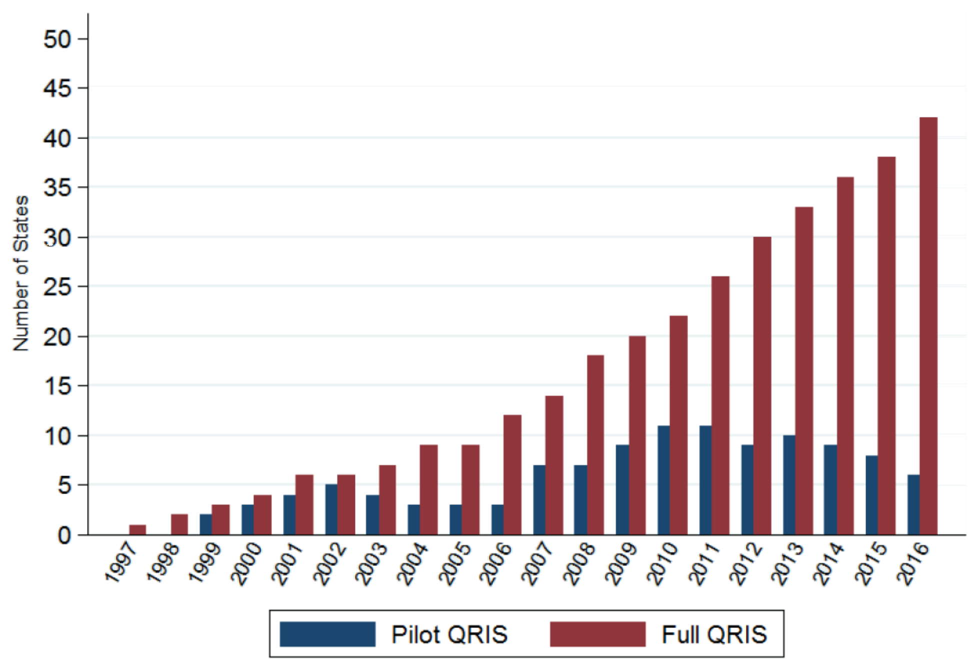


Figure G.1: Number of States with a QRIS Implemented

Notes: Figure taken from Herbst (2018)