Spring 2023

Simulating Jobs Created by the New York Universal Child Care Act

Brandon Istenes

Simulating Jobs Created by the New York Universal Child Care Act

Masters Paper Submitted to
Levy Economics Institute of Bard College

by Brandon Istenes

Annandale-on-Hudson, New York
May 2023

Acknowledgements

I would like to express my deepest gratitude to my advisor for this paper, Fernando Rios-Avila, who consistently provided timely and insightful advice and encouragement, and helped shape this project. One of this paper's contributions to the Levy Institute Micro Model is also due to him, that of fuzzing labor force participation probit to improve dispersion of job-takers. I am also deeply indebted to Thomas Masterson, who not only shared some of his code for the LIMM with me, but contributed a great deal of time and patience to helping me understand and adapt the model, and provided helpful suggestions on an early draft of this paper. I would also like to express my thanks to Emily Eisner for introducing me to this pressing topic, and to both her and James Ostaszewski for their help orienting my work with respect to it.

PLAGIARISM STATEMENT

I have written this project using in my own words and ideas, except otherwise indicated. I have subsequently attributed each word, idea, figure and table which is not my own to their respective authors. I am aware that paraphrasing is plagiarism unless the source is duly acknowledged. I understand that the incorporation of material from other works without acknowledgment will be treated as plagiarism. I have read and understand the Levy Economics Institute of Bard College statement on plagiarism and academic honesty as well as the relevant pages in the Student Handbook.

Brandon Istenes

ABSTRACT: The Universal Child Care Act recently proposed in the New York State Senate would begin the creation of a universal child care system in New York State. This would involve a large scale-up of child care service supply, precipitating a large increase in employment in the child care sector, while increasing wages for jobs in that sector. Who will benefit from these new jobs and wages? We use the Levy Institute Micro Model (LIMM) to simulate the distribution of these new jobs and wages to the population of New York State. Two econometric contributions are made to the LIMM which improve dispersion and result in allocations which are more representative of predicted likelihood distributions. The distribution of jobs and wages is found to be highly income-progressive, making it an effective pro-equity and anti-poverty measure. The distribution of jobs and wages favors women, especially women of color, across the state.

INTRODUCTION

The Universal Child Care Act recently proposed in the New York State Senate is addressed to pressing needs faced by families across the state. The New York Child Care Availability Task Force, chartered in 2018, published their final report in 2021 (NY OCFS 2021). Their findings highlight that the child care sector in New York is plagued by poverty wages, high turnover, "deserts" of little-to-no availability, and some of the highest user costs in the country and the world. The proposed Act seeks to address all of these problems through a combination of supply-side and demand-side interventions aimed at producing a universal child care system in New York State.

Child care in New York is some of the most expensive in the United States, and child care in the United States is some of the least affordable in the world (Gromada and Richardson 2021). It is a generally expensive state located in a country where there is less government support for child care services than almost any other developed country (Miller 2021). Means-tested subsidies are available for families under 300 percent of the poverty line (Moore 2022; NY OCFS n.d.). These subsidies constitute New York's demand-side support. It is a means-tested program, as opposed to a universal one. The inefficiency and non-take-up problems associated with means testing as a practice are well-known (Townsend and Gordon 2022, chap. 7); it is unlikely that these problems do not pertain to child care in New York.

New York also provides supply-side support to the child care sector. During the pandemic, a program of stabilization grants was "extraordinarily successful" at maintaining child care capacity (Nabozny 2022). There is an ongoing grant program allocating $100M to expand child care service capacity in child care deserts (NY OCFS n.d.). The New York Universal Child Care Act would expand these supply-side programs, add focused supply-side interventions to increase wages, and expand demand-side coverage to include all New Yorkers. Nevertheless, availability of child care is limited in much of the state. At the same time, child care workers struggle with extremely low wages, leading to poverty and dependency on government support. Further, limited funding means that only 5 percent of child care providers participate in QUALITYstarsNY, the state's quality rating and improvement system (NY OCFS 2021).

High-quality child care makes a significant improvement to the educational and life outcomes of children, especially those from disadvantaged backgrounds. Subsidizing child care encourages gender equity in employment and womens' lifetime earnings, and benefits from the economic activity of the sector ripple through the rest of the economy. The employment and income provided by child care sector expansion also tends to be highly income-progressive. Robust accounting of the costs versus the direct, indirect, and long-term state revenues induced by such programs tends to demonstrate net-positive fiscal impact (De Henau 2022; Fortin, Godbout, and St-Cerny 2012).

There are several mechanisms by which government spending on child care can reduce income inequality. The most important ones are the reduction in child care costs to families, increased income via maternal employment, and increased employment and wages in the child care sector. This paper is addressed to the last of these. We investigate how new employment and wages in the child care sector will be distributed across demographic and income groups if the New York Universal Child Care Act passes and comes into effect. A rough estimate of the number and types of new jobs that would be needed is produced, and a simulation is used to investigate the distribution of these new jobs among the population.

BACKGROUND

Child care in New York is some of the most expensive in the United States. The average annual cost of child care ranges from $10,140 for Family-Based care of a 4 year old (in 2018) to $15,028 for Center Based Infant Care (in 2017) (NY OCFS 2021). Subsidies are available for families under 300 percent of the Federal Poverty Line. Child care is fully subsidized for families under the poverty line ($24,860 for a family of three in 2023). Families with incomes between 100 percent and 300 percent of the poverty line must contribute copay for child care, which is a percentage of their income in excess of the poverty line that varies by county from 10-35 percent. Two parents both working full time for $17/hour with one child are not eligible for subsidies based on their incomes.

Most of New York State has substantially fewer child care slots than children. 64 percent of New Yorkers live in "child care deserts," which are areas with more than three times as many children as child care slots (NY OCFS 2021). There is an extent to which this makes sense in terms of supply and demand—very high costs mean that there is only demand for child care for a small percentage of children. However, geographic differences in the proportion of children to child care slots suggests that capacity is supra-linear with respect to demand ("Mapping the Gap™ in New York" n.d.). It seems likely that the fixed costs of providing child care services cause a deficiency of supply with respect to existing demand. Further, the sector has contracted dramatically in the pandemic. The share of children under 4 years old receiving child care fell from 34 percent in 2019 to 21 percent in 2020 (Nabozny 2022). Supply has luckily been relatively inelastic to the downward change in demand. The number of child care programs operating fell by only 8 percent from January 2020 to July 2022, and the amount of child care capacity fell by only 2 percent. However, even in good times, capacity changes suggest problems in the market provisioning of child care. From the NY Child Care Availability Task Force Final Report, "From 2011 to 2017, only the top 20 percent wealthiest communities saw an increase in infant/toddler capacity per 100 children ages 0-5."

The contraction in demand has worsened an already difficult situation with the employment of child care workers. The average annual wage of child care workers statewide in 2022 was $31,885 ("Occupational Wages" n.d.). It was estimated in 2016 that 65 percent of child care providers nationally are eligible for social safety net programs such as food stamps or Medicaid (Paquette 2016). In New York, aggregating data from 2016-2021, 40 percent of child care workers are actually receiving either food stamps or insured through Medicaid, compared to only 19 percent of the employed population generally (author's calculations from ACS). Between 2019 and 2021 the number of child care workers in the state declined by 43 percent (*ibid*), though by August 2022 it was estimated that the workforce was back up to 10 percent under pre-pandemic levels (Nabozny 2022). This despite the fact that wages from 2018 to 2022 have only risen by 16 percent, which is less than the CPI inflation rate of 18.8 percent over the same period (NY OCFS 2021; "Inflation Calculator" n.d.) (author's calculations). Under these conditions, staff retention is understandably difficult. To quote the memorandum for the New York Universal Child Care Act,

> Many professionals either do not enter the field at all, or are forced to leave it for higher paying jobs at places like fast food chains or telemarketing companies. This is not only an injustice, but also a tremendous vulnerability for our child care system because the understaffing problem leads to further shortages of available child care. (Brisport n.d.)

In the absence of affordable child care services, the burden of child care falls on families, and within families this responsibility falls overwhelmingly to women. Having little or no discretionary time is known as time poverty. Women's time poverty leads to fewer economic opportunities and adverse health consequences (Hyde, Greene, and Darmstadt 2020). This results in lower female labor force participation than in places with universal child care (Fortin 2017a). Without universal child care, the exorbitant cost of child care services has a coercive influence on womens' time use decisions, in many cases foreclosing any other option than full-time parenting.

Existing Programs

There are many subsidized child care programs around the world, with the US providing among the least parenting support of any wealthy country (Miller 2021). Toward the top of the UNICEF rankings for child care affordability are Italy, Germany, and Chile (Gromada and Richardson 2021).

Child care in Italy is subsidized by the government for children ages 0-6. Care for children ages 0-3 varies depending on the region and type of facility, but the average cost to the family is 301 euros per child (Bulgarelli n.d.). This is offset for many parents by an 11-month stipend of 300 euros per month beginning at the end of the paid maternity leave (SPLASH-db.eu 2014). Preschool is almost universally available from age three, and fees are generally much lower (Bulgarelli n.d.). Once all subsidies are taken into account, the average net cost of child care in Italy is approximately zero (OCED 2022).

Child care in Germany is also heavily subsidized by the federal government. States and communities organize the provision of child care, much of which is private, but very little of which is for-profit (Wrohlich 2005). Average net child care costs are also close to zero (OCED 2022).

Child care in Chile is also subsidized by the government. Publicly subsidized care is given by priority to low-income families. Children ages 4-5 are legally entitled to a place in preschool (OCED 2016). Net private expenditure on child care is zero (Gromada and Richardson 2021).

While the US in general has among the least public investment in child care (Miller 2021), there have been a number of programs providing some amount of support or subsidy for its provision. These include the Head Start and Early Head Start programs, and the Child Care & Development Fund. During World War 2 the US briefly ran a program providing universal child care services. The state of New Mexico last year (2022) amended the state constitution to guarantee early childhood education as a right and establish a steady stream of funding for child care services over the long run (Covert 2022).

At the state level, the program most comparable to the one being considered in the New York State legislature is the universal child care program in Quebec. This program started in 1997, making child care universally available for a low standardized rate—currently less than CAN$9. Since the inception of the program, women's labor force participation in Quebec has risen much faster than in the rest of Canada. When accounting for this increase, as well as changes to child care facility subsidies, tax subsidies, own-source government revenues, and family transfers due to this increase, the program is found to more than pay for itself. In 2008, it cost Quebec $1.2 billion, and increased Quebec's revenue by $1.4 billion, as well as increasing federal revenue by $673 million. This estimate does not take into account the indirect and induced employment and income created by the program in other sectors, which further increases both revenues and employment impact. The program is extremely popular among parents (Québec 2022a). The child care sector in Quebec was also much less affected by the Covid pandemic than that in the US (Hurley 2021).

The Quebecois program has faced an onslaught of criticism from conservative think tanks and right-wing economists arguing that it results in worse outcomes for children[1]. The basis for this comes from aggregating different types of care facilities. Non-profit child care centers called *centres de la petite enfance* (CPEs) serve 35 percent of all children in Quebec and have invariably been found to have positive cognitive, health, and behavioral impacts on children (Fortin 2017b; Geoffroy et al. 2010; Herba et al. 2013; I. Laurin et al. 2015; J. C. Laurin et al. 2015)

On the other hand, 17 percent of children attend private for-profit care centers called *garderies*, which generally provide low-quality care (Fortin 2017b; Lavoie, Gingras, and Audet 2015). Studies which find negative impacts of Quebec's universal child care on developmental outcomes do so only by aggregating across different types of care providers, ignoring the differential impacts of quality of care on child development. Doing so opens the door for bad-faith criticisms of universal child care or of professional child care services as a whole, rather than advocating for changes to the program that would provide better-quality care for children currently receiving substandard care.

The New York Universal Child Care Act

The New York State Senate is considering The Universal Child Care Act, introduced by Senator Brisport. The act would empower the preexisting Child Care Availability Task force, and a new office of early education, to guide New York toward a system of free and universal child care[2]. This would be accomplished through a combination of supply-side and demand-side interventions. The key goals are that the child care system be free to users, have sufficient capacity to meet the increased demand, and that child care workers should have wage parity with public school teachers. State reimbursements for care services will be increased to cover the costs to consumers and increased wages. The intention is that "New York State restructures its

[1] e.g. Lefebvre, Pierre, and Merrigan, Philip. 2008. "Child-Care Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada," Journal of Labor Economics, Vol. 26, No. 3, pp. 519–548; Baker, Michael, Gruber, Jonathan, and Kevin Milligan. 2008. "Universal Child care, Maternal Labor Supply, and Family Well-Being," Journal of Political Economy, Vol. 116, No. 4, pp. 709–745; Kottelenberg, Michael, and Lehrer, Steven. 2013. "New Evidence on the Impacts of Access to and Attending Universal Child-Care in Canada," Canadian Public Policy, Vol. 39, No. 2, pp. 263–285); Churchill, Aaron. 2019. "A cautionary tale: Universal childcare in Quebec." Fordham Institute.

[2] The Universal Child Care Act, S.7595, New York State Senate (2021)

economy to reflect the true value of this important work." Funds are provided to facilitate this transition both with demand-side subsidies and supply-side grants and other funds. One of these funds is a $1 billion workforce stabilization fund, which would supplement child care worker wages while this restructuring is underway, enabling the sector to hire and retain workers and provide living wages in the near term. Other funds go to increased subsidies for the provision of care services, grants to expand or start new child care centers, and investment in child care infrastructure.

LITERATURE REVIEW

Child care and child care subsidies are extensively studied in economics literature. Key questions that the literature typically seeks to address are the impact on female labor force participation, inequality and family finances, time use for families, government fiscal stance, inter-sectoral multiplier effects, and macroeconomic indicators such as GDP. The majority of studies are ex-post, with a more limited set of literature aimed at simulating ex-ante probable effects of policies under consideration. A variety of econometric techniques are used across the literature.

Among the features of child care programs most-studied by economists is the effect on female and maternal labor force participation (FLFP). Extensive reviews of methodologies for investigating FLFP can be found in Brewer and Paull (2004) and Kalb (2009). In brief, maternal labor force participation was predominantly modeled using a continuous utility maximization framework up until the 1990s. Since then, the most common methodology for modeling maternal labor force participation has been using utility maximization in a discrete decision framework.

A novel approach to investigating ex-ante the impact of investment in child care is taken by Aran et al. (2018). Their model integrates supply-side and demand-side models, and focuses on expanding child care service capacity. While this model is not used in this paper, it constitutes a major novel contribution to the child care policy simulation literature, and is relevant to themes of particular concern in New York State.

The present paper is concerned with assessing ex-ante the impact of increased employment and wages in the child care sector. Most research which does this uses the Levy Institute Micro Model (LIMM), which is used in this paper. An excellent investigation of methodological developments related to LIMM is provided in De Henau and Himmelweit (2020). A brief overview of some of the applications of LIMM and other care sector investment models will be provided here.

Capacity Simulation

A novel approach to investigating ex-ante the impact of investment in child care is taken by Aran et al. (2018). Their model integrates supply-side and demand-side models, and focuses on expanding child care service capacity. Three policy alternatives were studied: investment grants to service providers, operational grants to service providers, and child care vouchers to families.

The supply-side model uses the cost and pricing structure of child care centers, including initial investment and operational costs, to calculate the net present value of each child care center, and from that calculate the probability that the investment to open the child care center would be made in the first place. Given that all of the data used to fit the model is from actually existing child care centers, there may be some methodological questions worth asking about fitting a "probability of investment" curve using exclusively positive cases. That aside, estimated changes to child care sector capacity are calculated by introducing the grants as shocks.

Those changes to child care sector capacity are then introduced as exogenous shocks to the demand-side model. The demand-side model calculates a propensity to attend preschool for each child, and then allocates new capacity to those that have a high propensity but are not yet enrolled. From this, the number of children expected to newly enroll is predicted, as well as distributions of demographic features of those children. These distributions show who benefits from the increased capacity and cost reductions. They then estimate fiscal impact, and the cost-effectiveness per child enrolled.

They find that supply-side grants are both the most cost-effective and pro-poor interventions; that is, government expenditures on the supply-side yielded larger capacity increases in general as well as larger enrollment increases for poor households than income targeted demand-side vouchers. In terms of fiscal impact, they find that investment grants pay for themselves many times over, while operational grants and vouchers end up with a positive net cost, at least in the short run.

Job Creation and Income Effects

Most policy simulations investigating job creation and income effects use the Levy Institute Micro Model (LIMM), which is used in this paper. Two notable exceptions are AK Europa (2013) and Henau et al. (2016).

AK Europa (2013) simulates the potential aggregate employment and fiscal impacts of providing start-up grants for new child care centers. They calculate the costs of starting and running a child care center, and the number of child care centers that would be needed to meet current excess demand. They multiply the two numbers together to obtain the recommended magnitude of a start-up grant program. From the number of new child care centers, three types of employment effects are calculated: direct effects due to hiring new child care workers, indirect effects via inter-sectoral multiplier effects, and effects due to increased maternal employment among families using child care. Fiscal impacts are then estimated. They estimate that even in the most pessimistic case the grant program would more than pay for itself in four years.

Henau et al. (2016) construct a comparative analysis of employment stimulus across seven countries with a focus on care sector investment. They use input-output analysis to compute employment multipliers, providing an estimate of the number and types of new jobs that would be created by the investment. They then assume that the distributions of characteristics of people who fill these new jobs are identical to the distributions of characteristics of the people that presently work in those jobs. They find that investment in care sectors has “slightly better” results on overall employment, and reduce gender gaps more effectively than “gender-neutral” investment strategies, while increasing male employment in equal measure.

The Levy Institute Micro Model (LIMM) was created in 2009 to simulate the employment and income impacts of the American Recovery and Reinvestment Act (Masterson 2023; Zacharias, Masterson, and Kim 2009), though it appears connected with a 2007 simulation project for South Africa undertaken by Antonopoulos (see Antonopoulos (2008)). It is described first by Antonopoulos et al. (2010), and then in greater technical detail by Masterson (2013). It has been used in many countries; here we will discuss Antonopoulos et al. (2010) about the United States, Kim et al. (2019) about Turkey, and De Henau (2022) about the United Kingdom.

Antonopoulos et al. (2010) create and describe the LIMM, and apply it to the United States to investigate the job creation potential of investment in the social care sector. Their model has two main parts: calculation of multipliers using input-output matrices and simulated allocation of jobs and income using microdata. Jobs are allocated to the unemployed. Their simulations find that investment in the social care sector produces almost twice as many jobs per dollar as investment in infrastructure. The allocation of jobs and income from investment in social care is also much more favorable to women, much more evenly distributed across racial groups, and highly progressive with respect to household income and level of education.

Kim, İlkkaracan, and Kaya (2019) extend this two-step model to examine the impacts of fiscal expenditure targeting supply-side expansion of early childhood and preschool education services in Turkey. They use the multipliers obtained from the input-output tables to estimate the cost of expanding child care services in Turkey to meet the OECD average enrollment rate, and the number of new jobs required. They estimate both the gender distribution of employment and the benefits and income provided by those jobs using the same allocation method as above. However, in this case, jobs are allocated not just to the unemployed, but also those who might plausibly enter the labor force. The propensity regressions used also include more independent variables. They find that expansion of the child care sector creates more than twice as many jobs as would be created by the same expenditure directed toward construction or direct cash transfers. In addition, these jobs tend to have higher pay, better benefits, and be permanent, though these features are probably more specific to the Turkish context, where construction jobs are frequently temporary or without contract. A substantial majority of these jobs are estimated to go to women. The distribution of jobs and new income by educational attainment and by income is less progressive; however, it is likely that these results (as well as the magnitude of the gender-differential impact) would be quite different in the US due to differences in the demographics of child care workers and the labor force generally. Despite this, the income-enhancement effect is highly progressive, with the average earnings in job recipient households in the bottom income quintile seeing their incomes increase by 58 percent, compared to a 8 percent increase for households in the top income quintile. The question of short-run fiscal sustainability is also addressed, leaving aside the question of long-run economic benefits from improved child care services. It is estimated that the government would recover 77 percent of expenditures on the program via increased government revenues in the short-run.

De Henau (2022) presents an ex-ante simulation of a hypothetical universal child care program for the United Kingdom. They estimate the annual public expenditure that would be required to provide universal care to children 6 months to 4.5 years, using a similar method to Kim, İlkkaracan, and Kaya (2019). They first estimate the full costs of a universal child care program, including scenarios involving higher qualification and pay for child care staff. They calculate costs required to increase capacity required for two different scenarios for level of uptake into the program. Direct, indirect, and induced employment effects are then calculated using input-output matrices. Allocation of these jobs follows the same method as above. They then calculate the change in maternal labor supply, tax revenue, and reduced family spending, differentiated by income group. These costs and revenues are compared to obtain the estimated fiscal impact of the program. In the short run the program is estimated to have a positive net cost between 0.3 percent and 0.8 percent of GDP. However, it is estimated that the increased average lifetime earnings of mothers offset the costs in the long run, and it is suggested that lifetime productivity gains for children could also contribute to covering the cost.

This paper is thus taking a well-trodden path in modeling child care investment by adapting the LIMM. At the same time, the LIMM is an evolving tool, with many moving parts that are selectively applied based on context, data availability, and other concerns. It is a natural choice for the present inquiry. Hopefully the contributions made to the LIMM here may give back some to the tradition on which this paper draws.

METHOD

The question addressed in this paper is how new jobs in the child care sector, and income from those jobs, are distributed across demographic and income groups. There are two steps to determining this. First, a rough estimate of the number and kinds of new jobs that would need to be created for occupations in the child care sector is calculated based on capacity shortfalls and comparison with Quebec's program. Then, given that distribution of job opportunities, one needs to estimate how those jobs are distributed across the population.

As discussed above, this paper will use the Levy Institute Micro Model (LIMM). However, most usages of the LIMM calculate employment multipliers from an input-output table in order to determine the total distribution of new jobs created. There was no input-output table for the state of New York available to this author. Thus only the direct employment effects are used, and only the half of the LIMM concerned with job allocation is used.

Creating New Jobs

The Act is directed at the provisioning of universal child care in New York State. Therefore the key question in determining the number and kinds of jobs that would be required is: what would be required to expand the capacity of the child care sector sufficiently to provide universal access? There are two ways in which we might define the level of capacity required. One is the "universal means everyone" approach, where capacity would be expanded sufficiently to provide professional child care services to all children in the state. The other is a comparative empirical approach, where we assume that capacity needs to be expanded to suffice for a reasonable prediction of utilization rates given universally accessible child care. Put another way, if subsidies provided were sufficient for any family that wants child care to afford it, what would the expected demand be? How many new child care slots do we need to create, and how many new jobs are needed to support that?

This question is answered by looking at our near neighbor, Quebec, which has universal child care; it is available to any family that wants it at a heavily subsidized daily rate. In Quebec 72 percent of children of preschool age attended child care on a regular basis, and 79 percent of all children have been enrolled in La Place 0-5, the state daycare registry (Québec 2022b).

Therefore, it is assumed that if child care was made universally accessible and affordable throughout New York State, slots would be required for about three of every four children.

The application of utilization rates from Quebec to the New York context is justifiable based not just on geographic proximity, but on demographic similarity in several important measures. The New York State FLFP rate has been roughly constant since 1996, hovering between 53.9 percent and 57.3 percent. This leveling out after the late 1990s is seen across the United States (U.S. Bureau of Labor Statistics 2022). This is approximately equal to the one that held in Quebec in 1996, 54.4 percent. Since the institution of universal child care in Quebec the FLFP rate has increased to 61.6 percent in 2023 (Statistics Canada 2023). The trend in the rest of Canada, which does not have universal child care, is similar to that in the United States, with FLFP flattening out after the late 1990s (Fortin 2017a). Further, family size is comparable in the two regions. In both places, 57 percent of married couples have children (Statistics Canada 2023) (author's calculations from ACS data for NYS). In both regions, the average number of children among families that have any children is 1.8 (Québec 2018) (author's calculations from ACS data for NYS).

The shortfall in child care slots is based on research by the New York Child Care Availability Task Force completed in 2021 (NY OCFS 2021). They estimate that there are approximately four children per slot statewide, and six children per slot when excluding New York City. By assuming that child populations are proportional to overall populations we calculate that the New York City proportion is 2.8 children per slot. Thus child care sector expansion is modeled separately for two geographies, New York City and the rest of New York State.

The number of new child care jobs required to meet the shortfall is obtained using the strong assumption that the number of people employed in each occupation is in constant ratio with the number of child care slots. Pools of new jobs are per-occupation and per-region (NYC and outside NYC). Thus the two pools of new jobs needed for a given occupation in the child care industry are obtained by multiplying the current number of workers in that occupation and region by the child-slot ratio (2.8 for NYC, 6 for outside NYC), and then multiplying by the proportion children that are expected to need slots, 0.75.

This assumption of constant slot-to-worker ratios is one of the places the model could benefit substantially from improvement. The structure of capacity changes in the child care sector are more complicated than that. This is clear from looking at the effect of the pandemic on the child care industry in New York State. From 2019 to 2020, the number of children enrolled in child care declined by almost one third (Nabozny 2022). This is similar to the job losses among child care workers, where from 2019 to 2021 there has been a 43 percent decline in the number of child care workers (a 37 percent decline in all employment in the child care industry) (author's calculations from ACS data). However, from January 2020 to July 2022, the net loss of child care programs was only 8.1 percent (Nabozny 2022), and from 2019 to 2022 the number of child care administrators declined by only about 5 percent (author's calculations from ACS data). This suggests that a significant amount of capacity increase and decrease is accommodated by hiring and firing child care workers within existing programs. It seems likely that the number of child care workers is approximately linear in proportion with the number of child care slots, while the number of child care administrators is approximately linear in proportion with the number of child care programs. However, this paper makes no attempt to model the opening of programs and structure of capacity increase, so a linear relationship from slots to jobs is used for all occupations. This is probably a better approximation than the pandemic data would suggest, since an expansion of capacity needed for universal child care would almost certainly be far in excess of the potential capacity of existing programs and therefore require opening many new programs. Still, it is likely that this model overestimates the number of administrator jobs that will need to be created. This is especially the case since the proportion of administrators to slots is based on data that includes the pandemic years, during which the administrator to slot ratio was especially high.

Distribution of New Jobs

An extension of the Levy Institute Micro Model developed at the Levy Economics Institute is used to simulate the hiring of workers from the unemployed population. Using this simulation, the distribution of new child care jobs across demographics and the impact of those jobs on incomes is examined.

Data used in the model comes from the ACS 5-year survey consisting of data from 2016 to 2021. It should be noted that this includes the COVID pandemic, which caused a substantial disruption in the child care sector. The nature and impact of the pandemic on the reliability of the model is described above in the discussion of new job creation. The model considers a person as having a child care job if they are employed in the child care industry with one of the following three occupations: Child care worker, Teaching assistant, and Education and Child care Administrator. The sample includes 2,611 people who have child care jobs according to this definition. A limitation of the data is that only 156 of these people are men, so estimates about men will be lower quality than those for women.

It should also be noted that the salary distributions of these occupations differ somewhat from the salary distributions for child care occupations identified in the New York Cost of Quality Study (Workman and Jessen-Howard 2019), and it is not obvious how to map the two sets of occupations together. People in the Administrator occupation in the ACS data report incomes comparable to Director positions in the Cost of Quality Report, though the distribution for New York City has a fat tail upward, resulting in a mean well in excess of the median. People in either the Child Care Worker or Teaching Assistant occupation in the ACS data report incomes consistent with the Floater or Teacher Assistant occupations.

The basic idea of the model is to examine the demographic characteristics of people who currently have child care jobs ("donors") and then assign new child care jobs to people who are eligible ("recipients") in order of their demographic similarity to the people who currently hold those kinds of jobs. The eligible population is all unemployed people between the ages of 18 and 74 who reported either looking for work, being available for work, or being in the labor force. The cutoff age of 74 is determined heuristically from the distribution of ages of people who presently work in the child care sector.

Two probit models are run, one for men and one for women, to identify the demographic features of people who work in the child care sector generally. Demographic features considered are age, race, marital status, and number of children under five. These probit models are used to predict the likelihood that people from the eligible population will get jobs in the child care sector. A normal distribution fuzz is applied to coefficients to provide a more realistic dispersion of new job takers. The variance of the fuzz distribution is determined heuristically from the standard deviation of the coefficients. The fuzzing of the coefficients is a new addition to the Levy Institute Micro Model. A multinomial logit model is then estimated to predict the probabilities of taking specific occupations within the child care sector based on demographic features.

Job offers are then simulated, and the interested population is determined. The interested population is the subset of the eligible population who is likely to take such a job if offered. A simulated job offer consists of a wage and a number of hours per week. To obtain job offers, first wages and hours are regressed against demographic features and the occupation likelihoods for current child care jobs holders, within region and sex cells. Then wages and hours are predicted from this regression for the eligible population. These predictions are used along with demographics, spouse demographics, and employment likelihood to produce job offers in regional cells using multiple imputation with hot decking. The Inverse Mills Ratio is used to correct for selection bias. Each person then decides whether they are interested in the offer based on whether the offered income is greater than 75 percent of what they made last year, with some random normal fuzz applied to improve dispersion.

The other contribution to the Levy Institute Micro Model in this study is a revision of the algorithm for assigning jobs once employment likelihood, occupation likelihood, and the interested population are determined. In order to ensure that both employment likelihood and occupation likelihoods are taken fully into account when allocating jobs, the joint probability of these likelihoods are used. That is, the two probabilities are multiplied together and normalized. This produces, for each person in the interested population, the probability that a new job in a given occupation will be allocated to that person. Jobs are then allocated by simulating draws from the interested population according to these probabilities, until all the jobs are allocated (or there are no more eligible and interested people to take them).

Due to Stata's `rdiscrete`[3] having a 10,000 row limit, draws from the interested population must be simulated by first drawing random samples of 10,000 from the interested population, and then

[3] https://www.stata.com/manuals13/m-5runiform.pdf

using `rdiscrete` to draw 1 percent of those samples based on the joint employment-occupation probabilities (so 100 probability-based draws per random sample of 10,000), until all the jobs are taken or there are no more people to take them. This process simulates drawing from the discrete probability distribution of the entire interested population.

A key goal of the Act is to increase child care worker incomes. Section 15 of the Act states that child care workers' salaries should be brought to parity with those earned by K-12 teachers. This is to be accomplished in the long run by a transition from market rate based reimbursement rates to "true cost of care" based reimbursement rates, where increased salaries are understood to be part of the true cost of care. In the short run, the Act calls for the creation of a Child Care Workforce Stabilization Fund which would supplement worker wages directly until such a transition is complete. Since the job allocation model used here is not in any way temporal, the question of where wage increases come from at the time of job allocation is left aside. A uniform wage increase factor is applied to all existing child care job incomes and all new job offers. This factor is based on the New York State Cost of Quality Child Care Study (Workman and Jessen-Howard 2019), which finds that wage parity with K-12 teachers would require wages to increase by 83 percent.

RESULTS

The results of this study suggest that the New York Universal Child Care Act would have a significant impact on job creation, income growth, and economic equity throughout the state. Four scenarios were simulated. Scenario A is for the provisioning of enough child care slots for 75 percent of the state's children, with a uniform wage increase of 83 percent, corresponding to the increase in wages to parity with public school teachers. In Scenario B wages are only increased by a uniform 36 percent, which is a level consistent with a $15 minimum wage based on the New York Cost of Quality Study. Scenario C investigates the provisioning of child care slots for 95 percent of the state's children, with the 83 percent wage increase. Scenario D attempts to model a progressive wage increase by increasing wages by 83 percent for child care workers and teachers' assistants but only by 50 percent for administrators.

The creation of 206,000 jobs in the child care sector would be necessary to expand child care services to provide capacity for 75 percent of the state's children, which is Quebec's level of demand for their program of universal child care. 87-88 percent of these new jobs would go to women. Meeting demand for 95 percent of the state's children would require the creation of 261,000 jobs, of which 83 percent would go to women. Scenarios A and D result in $7.5 billion in new wages. Lower wages in Scenario B result in only $5.7 billion in new wages. More jobs in Scenario C results in $9.8 billion in new wages. Per the model design, these benefits only account for the new jobs created directly in the child care sector. It is likely that accounting for induced employment and income would substantially increase this estimate of the economic output of the state.

A large share of the economic benefits of new jobs would go to economically disadvantaged demographic groups. According to the simulation, in New York City, 25 percent of jobs would go to Black non-Hispanic women, 42 percent to Hispanic women, and 80 percent to women of color in general. Statewide, 45 percent of jobs would go to women of color. These proportions are roughly the same across scenarios. Hispanic women receive a much larger share of jobs than the proportion of the eligible population that is Hispanic, while Black women receive a share of the jobs approximately equal to the proportion of the eligible population which is Black.

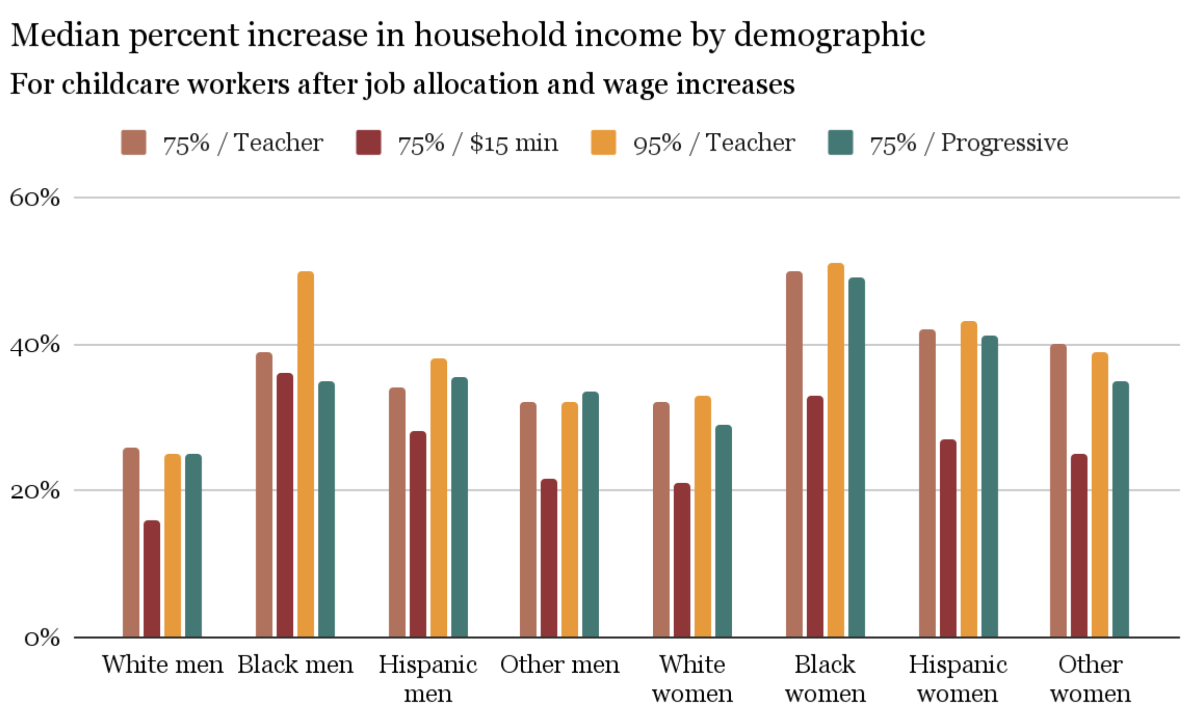
For those that work in the child care sector once new jobs are allocated, those that benefit the most in terms of their household incomes are Black women, and to a lesser degree, Black men, Hispanic men and other non-white women. Black men are the only demographic group for whom the number of jobs created has a significant impact on the median increase in their household income. This seems to happen because when more jobs are available, though a similar proportion of those jobs go to Black men, more of them go to Black men with lower household incomes. For all demographics except Black men, the wage increase has a large impact on the household incomes of people working in child care.

Table 1: Recipients of New Jobs in the Child Care Sector, Scenario A (75% / Teacher Wages)

NYC					
	White	**Black**	**Hispanic**	**Other**	**Total**
Male	2.51%	2.41%	3.67%	1.32%	9.92%
Female	17.10%	22.96%	37.60%	12.43%	90.08%
Total	19.60%	25.37%	41.28%	13.75%	100.00%
Reference pop	26%	25%	16%	33%	100%

Non-NYC					
	White	**Black**	**Hispanic**	**Other**	**Total**
Male	8.69%	2.14%	2.35%	0.94%	14.12%
Female	55.87%	10.51%	13.98%	5.52%	85.88%
Total	64.57%	12.65%	16.33%	6.45%	100.00%
Reference pop	68%	12%	7%	13%	100%

Figure 1: Median Percent Increase in Household Income by Demographic



The implementation of universal child care would have a significant impact on poverty and equity. Currently, 13 percent of people working in the child care sector in New York State are in poverty. Only 5 percent of the employed population in New York State generally is in poverty, so poverty rates among child care workers are much higher than average. In all scenarios, 28 percent of the new jobs would go to households under the poverty line, even though these households account for only 13 percent of the population. The median increase in household income among child care workers (both old and new) under the poverty line is estimated to be 135 percent, or about $27,000, in Scenarios A, C, and D; it is 100 percent or $21,000 in Scenario B. We calculate a proxy measure of poverty from the data as

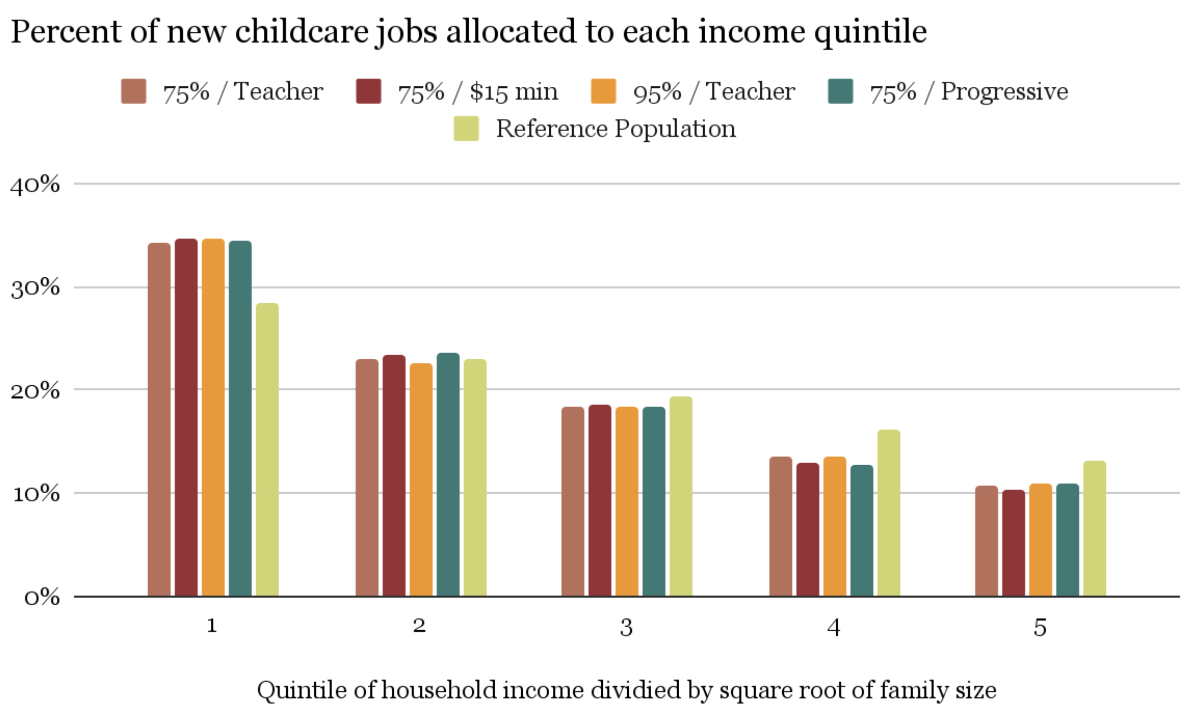
$$poverty = a \times householdIncome / familySize,$$

where a is a normalization constant selected to make the distribution of *poverty* similar to the poverty variable provided by ACS. We choose the value $a = 1/168$. We use this proxy poverty to estimate the number of people whose household incomes cross the 100 percent poverty line because of this act. We also use it as the measure of income stratification for quantile analysis of impact. In Simulations A and D, the number of people brought out of poverty by new jobs and improved wages in child care is about 45,000. In Simulation B, where the wage increase is smaller, 41,000 people are brought out of poverty. In Simulation C, where more jobs are created, 54,000 people are brought out of poverty.

Of course, one would hope that the number of people working in child care and living in poverty would be zero if this act passes. In the simulation, in Scenarios A, C, and D, the poverty rate for child care workers is reduced to 6.3 percent; in Scenario B it is reduced to 8.1 percent. The remaining poverty is likely due to the extrapolation of current patterns of part-time work and part-year unemployment. This may not be realistic; with higher effective demand for child care slots and higher demand for workers it seems likely that a larger share of child care workers would be working year-round and full-time, modulo structural time constraints for workers caring for school-age children. This would be expected to further reduce the poverty rate for child care workers.

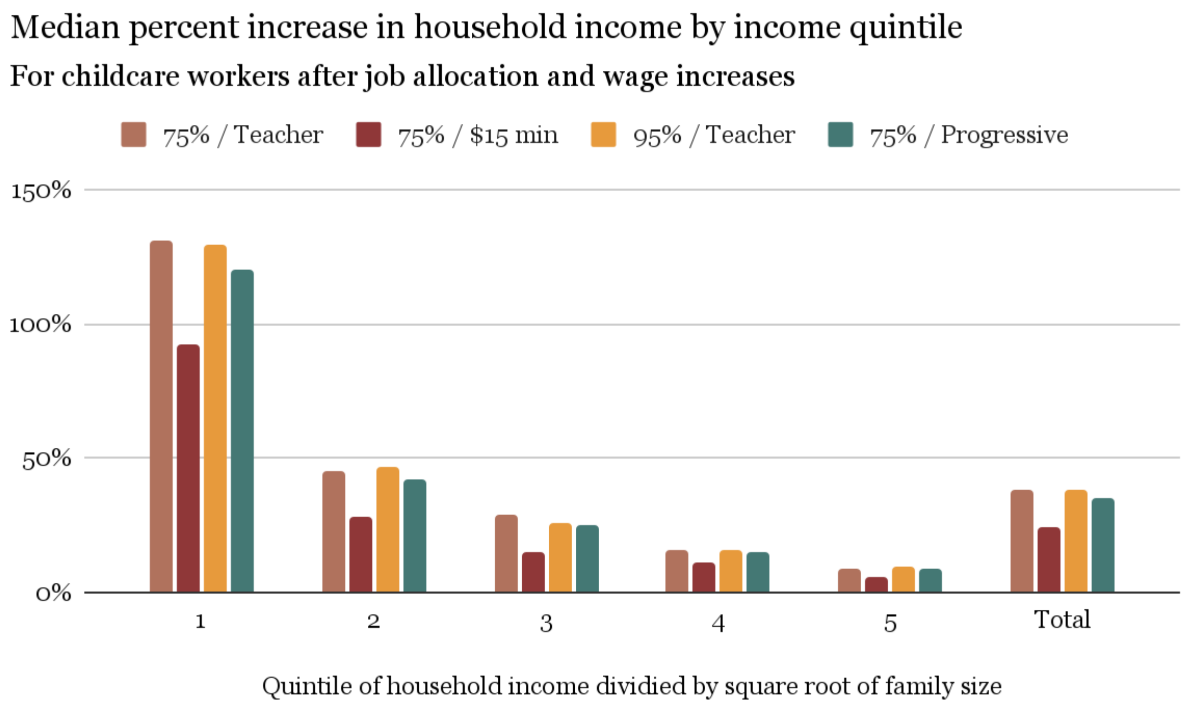
Taking our proxy for poverty as a measure of income, we analyze the quantile distribution of jobs and income. Distribution of jobs and income is highly progressive with respect to our income measure. Proportional distribution of jobs is roughly the same for all scenarios: 35 percent of jobs to the bottom income quintile, 23 percent to the next, and 18 percent to the middle income quintile.

Figure 2: Percent of New Child Care Jobs Allocated to Each Income Quintile



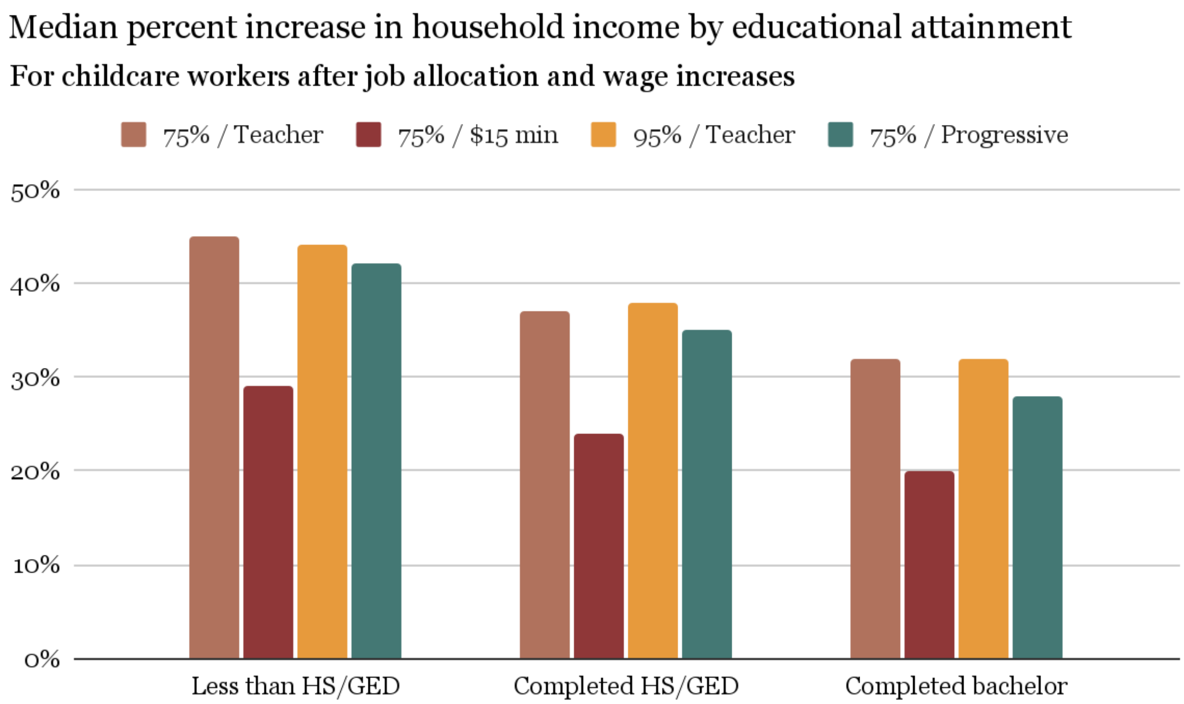
Median increase in earned income is roughly flat across income quintiles, and lower in Scenario B than the other scenarios. As a percentage increase of household income, the new incomes have a highly progressive differential impact, which is shown in Figure 2. Note that the percentage increases of household incomes refer to the household income itself, not the poverty proxy measure which is scaled by family size.

Figure 3: Median Percent Increase in Household Income by Income Quintile



Finally, we examine the distribution of new income with respect to educational attainment. In all four simulations, jobs are distributed across educational attainment with approximately the same proportions: 9 percent for people with less than high school education, 68 percent for people with high school equivalent education, and 23 percent for people with at least a bachelor's degree. Simulated wages increase slightly as the level of education increases, but by significantly smaller increments than current household income, with the result that households that benefit the most in relative terms are those with less education.

Figure 4: Median Percent Increase in Household Income by Educational Attainment



These results highlight the importance of investing in the child care sector as a means of promoting economic growth and equity in the state.

CONCLUSIONS

The simulation demonstrates that employment from the New York Universal Child Care Act would provide substantial differential benefits to Black women, Hispanic women, women of color, and women in general across the state. The distribution of new jobs and income from the New York Universal Child Care Act would be highly income-progressive, benefiting through employment some of the poorest households in New York State. It is an effective poverty reduction measure.

Methodologically, this paper makes two key contributions to the Levy Institute Micro Model. First, the introduction of coefficient fuzzing to the employment likelihood model provides more realistic dispersion. The more important contribution is the simultaneous use of employment and occupation likelihoods as the probability of drawing a job recipient from the pool. This ensures that neither likelihood goes to waste and that adequate dispersion across both is achieved, preventing potentially serious biasing problems.

This paper has focused on only one aspect of the act: the impact of employment it would create directly in the child care sector. There are many other aspects of the act that should be considered, even if we leave aside more philosophical questions about what it means to establish child care as a right, and the social value of providing child care to families regardless of income. It would be worthwhile to draw more on comparisons with other universal child care programs. It is worth considering, for example, ex-post research about social outcomes, educational attainment, and gender equity with respect to what might be possible in New York.

A number of other questions might fruitfully be approached using econometric ex-ante simulation methods. First, how do we expect the child care subsidies to affect inequality and poverty? Affordability for families is a key problem. This paper has addressed the impacts of new jobs in child care, but not the impacts for families, the users of child care services.

Another key question is how and how much capacity will expand. How much new child care capacity do we expect to be induced through supply-side grants and demand-side subsidies? The Act includes both, with separate budget line-items. Knowing how each of these tools impact the expansion of child care services will help the state to use funds strategically and help ensure the provision of sufficient high-quality services to meet New York families' needs. Further, the provisioning of services in child care deserts is one of the key motivations of the act. If we consider the geographic distribution of care, to what extent do we expect these provisions to eliminate the child care deserts in New York State? These questions might be approached using the technique in the Aran paper. Geographically-specific costing information could be used as a simple approach for estimating the differential effects around the state. A more nuanced approach might also take into account existing child care capacity density and populations of children in each region.

We might also be able to make estimates about budget utilization. How do we expect the funding allocated by the Act to be spent out? Will it be sufficient to achieve the goals of the Act? A dynamic version of the Aran model could help address the part of the question concerned with supply-side and demand-side subsidies, though the transition from market-based subsidies and wages to ones based on cost of quality care will require some novel work.

Finally, what do we expect to be the macroeconomic impacts of this act? What additional employment and income outside the child care sector do we expect that it would create? What do we expect to be the fiscal impact, accounting for new employment, new tax revenues, reduced dependence on government aid such as food stamps, and higher economic output? A robust accounting should include both indirect and induced effects on employment and income, as well as increased lifetime maternal earnings.

We must also remember to look beyond the econometric, at the big picture. As De Henau and Himmelweit (2020) write, "...while it is important to model and analyze the indirect benefits of investment, they should not be allowed to overshadow the main reason for considering investing public money in social infrastructure: the direct benefits that it brings for people in being healthier, better cared for, and better educated."

The existing body of research into these questions with regards to child care investment and subsidy programs elsewhere offers much reason to be optimistic. Still, focused engagement with data specific to New York can help show that these results are likely to pertain here, and may help guide the implementation of a successful program going forward.

REFERENCES

AK Europa. 2013. "Economic and Fiscal Effects of Improv Childcare in Austria." Brussels, Belgium: AK EUROPA. https://eige.europa.eu/sites/default/files/documents/main_report_en_304.pdf.

Antonopoulos, Rania. 2008. "POLICY BRIEF: Case-Study on South Africa - Impact of Employment Guarantee Programmes on Gender Equality and Pro-Poor Economic Development." https://www.levyinstitute.org/pubs/UNDP-Levy/South_Africa/Policy_Brief_EPWP_South_Africa.pdf.

Antonopoulos, Rania, Kijong Kim, Thomas Masterson, and Ajit Zacharias. 2010. "Investing in Care: A Strategy for Effective and Equitable Job Creation." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1659726.

Aran, Meltem A., Ana Maria Munoz-Boudet, and Nazli Aktakke. 2018. "Building an Ex-Ante Simulation Model for Estimating the Capacity Impact, Benefit Incidence, and Cost Effectiveness of Child Care Subsidies in Turkey." *International Journal of Child Care and Education Policy* 12 (1): 15. https://doi.org/10.1186/s40723-018-0052-3.

Brewer, Mike, and Gillian Paull. 2004. *Families and Children Strategic Analysis Programme (FACSAP): Reviewing Approaches to Understanding the Link between Childcare Use and Mothers' Employment*. London: Department for Work and Pensions.

Brisport, Jabari. n.d. "New York State Senate Memorandum in Support of Legislation: The Universal Child Care Act." The New York State Senate. Accessed May 3, 2023. https://www.nysenate.gov/sites/default/files/bill_memo_-_universal_child_care.pdf.

Bulgarelli, Daniela. n.d. "Quality of Employment in Childcare - Country Report: Italy." European Public Service Union. Accessed May 3, 2023. https://www.epsu.org/sites/default/files/article/files/Country%20report%20Italy%20childcare.pdf.

Covert, Bryce. 2022. "New Mexico Is the First State to Guarantee a Right to Early Childhood Education. Universal Child Care Could Come Next." *Early Learning Nation* (blog). November 9, 2022. https://earlylearningnation.com/2022/11/new-mexico-is-the-first-state-to-guarantee-a-right-to-early-childhood-education-universal-child-care-could-come-next/.

De Henau, Jerome. 2022. "Simulating Employment and Fiscal Effects of Public Investment in High-Quality Universal Childcare in the UK." *International Journal of Child Care and Education Policy* 16 (1): 3. https://doi.org/10.1186/s40723-022-00096-y.

De Henau, Jerome, and Susan Himmelweit. 2020. "Developing a Macro-Micro Model for Analyzing Gender Impacts of Public Policy." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3666317.

Fortin, Pierre. 2017a. "What Have Been the Effects of Quebec's Universal Childcare System on Women's Economic Security?," March.

———. 2017b. "Quebec's Childcare Program at 20." *Inroads* (blog). November 15, 2017. https://inroadsjournal.ca/quebecs-childcare-program-20-2/.

Fortin, Pierre, Luc Godbout, and Suzie St-Cerny. 2012. "Impact of Quebec's Universal Low Fee Childcare Program on Female Labour Force Participation, Domestic Income, and Government Budgets." https://www.oise.utoronto.ca/atkinson/UserFiles/File/News/Fortin-Godbout-St_Cerny_eng.pdf.

Geoffroy, Marie-Claude, Sylvana M. Côté, Charles-Édouard Giguère, Ginette Dionne, Philip David Zelazo, Richard E. Tremblay, Michel Boivin, and Jean R. Séguin. 2010. "Closing the Gap in Academic Readiness and Achievement: The Role of Early Childcare." *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 51 (12): 1359–67. https://doi.org/10.1111/j.1469-7610.2010.02316.x.

Gromada, Anna, and Dominic Richardson. 2021. "Where Do Rich Countries Stand on Childcare?" UNICEF Office of Research – Innocenti. https://www.unicef-irc.org/publications/pdf/where-do-rich-countries-stand-on-childcare.pdf.

Henau, J., S. Himmelweit, Zofia Lapniewska, and D. Perrons. 2016. "Investing in the Care Economy. A Gender Analysis of Employment Stimulus in Seven OECD Countries." In . https://www.semanticscholar.org/paper/Investing-in-the-Care-Economy.-A-gender-analysis-of-Henau-Himmelweit/79660a6f011dd8b137ff939fa5214173cc2320c9.

Herba, Catherine M., Richard E. Tremblay, Michel Boivin, Xuecheng Liu, Chantal Mongeau, Jean R. Séguin, and Sylvana M. Côté. 2013. "Maternal Depressive Symptoms and Children's Emotional Problems: Can Early Child Care Help Children of Depressed Mothers?" *JAMA Psychiatry* 70 (8): 830–38. https://doi.org/10.1001/jamapsychiatry.2013.1361.

Hurley, Kendra. 2021. "In Quebec, Child Care Is Infrastructure." *Bloomberg.Com*, April 29, 2021. https://www.bloomberg.com/news/articles/2021-04-29/lessons-from-quebec-on-universal-child-care.

Hyde, Elizabeth, Margaret E Greene, and Gary L Darmstadt. 2020. "Time Poverty: Obstacle to Women's Human Rights, Health and Sustainable Development." *Journal of Global Health* 10 (2): 020313. https://doi.org/10.7189/jogh.10.020313.

"Inflation Calculator." n.d. Accessed May 3, 2023. https://www.omnicalculator.com/finance/inflation.

Kalb, Guyonne. 2009. "Children, Labour Supply and Child Care: Challenges for Empirical Analysis." *Australian Economic Review* 42 (3): 276–99. https://doi.org/10.1111/j.1467-8462.2009.00545.x.

Kim, Kijong, İpek İlkkaracan, and Tolga Kaya. 2019. “Public Investment in Care Services in Turkey: Promoting Employment & Gender Inclusive Growth.” *Journal of Policy Modeling* 41 (6): 1210–29. https://doi.org/10.1016/j.jpolmod.2019.05.002.

Laurin, Isabelle, Danielle Guay, Michel Fournier, Nathalie Bigras, and Anabel Solis. 2015. “Attendance in Educational Preschool Services: A Protective Factor in the Development of Children from Low-Income Families?” *Canadian Journal of Public Health* 106 (S7): eS14–20. https://doi.org/10.17269/CJPH.106.4825.

Laurin, Julie C., Marie-Claude Geoffroy, Michel Boivin, Christa Japel, Marie-France Raynault, Richard E. Tremblay, and Sylvana M. Côté. 2015. “Child Care Services, Socioeconomic Inequalities, and Academic Performance.” *Pediatrics* 136 (6): 1112–24. https://doi.org/10.1542/peds.2015-0419.

Lavoie, Amélie, Lucie Gingras, and Nathalie Audet. 2015. “La qualité éducative dans les garderies non subventionnées. Faits saillants, Enquête québécoise sur la qualité des services de garde éducatifs 2014.” Québec, Institut de la statistique du Québec. https://statistique.quebec.ca/fr/fichier/la-qualite-educative-dans-les-garderies-non-subventionnees-faits-saillants.pdf.

“Mapping the Gap™ in New York.” n.d. ChildCare Aware of America. Accessed May 3, 2023. https://ccaoa.maps.arcgis.com/apps/MapSeries/index.html?appid=a7cace10038a41f283dd06705a2e612e.

Masterson, Thomas. 2013. “Quality of Statistical Match and Simulations Used in the Estimation of the Levy Institute Measure of Time and Consumption Poverty (LIMTCP) for Turkey in 2006.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2296203.

———. 2023. Research Methods II LectureLecture.

Miller, Claire Cain. 2021. “How Other Nations Pay for Child Care. The U.S. Is an Outlier.” *The New York Times*, October 6, 2021, sec. The Upshot. https://www.nytimes.com/2021/10/06/upshot/child-care-biden.html.

Moore, Kathleen. 2022. “Make $69,090? Your Kids Can Get Child Care Subsidies.” Times Union. August 3, 2022. https://www.timesunion.com/news/article/Childcare-subsidies-now-available-for-17345901.php.

Nabozny, Pete. 2022. “Rebuilding Our Future: Child Care Closures in New York During the Pandemic.” The Children’s Agenda. August 2022. https://thechildrensagenda.org/wp-content/uploads/2022/08/Child-Care-Closures-Report-FINAL.pdf.

NY OCFS. 2021. “Supporting Families, Employers & New York’s Future: An Action Plan for a Strong and Equitable Child Care System - Final Report of the Child Care Availability Task Force.” New York State Office of Children and Family Services. https://ocfs.ny.gov/reports/childcare/Child-Care-Availability-Task-Force-Report.pdf.

———. n.d. “Help Paying for Child Care.” New York State Office of Children and Family Services - Division of Child Care Services. Accessed May 3, 2023a. https://ocfs.ny.gov/programs/childcare/ccap/help.php.

———. n.d. “Invest in New York - Child Care Deserts Grant for New Providers.” New York State Office of Children and Family Services - Division of Child Care Services. Accessed May 3, 2023b. https://ocfs.ny.gov/programs/childcare/deserts/.

“Occupational Wages.” n.d. Department of Labor. Accessed May 3, 2023. https://dol.ny.gov/occupational-wages-0.

OCED. 2016. “Starting Strong IV: Early Childhood Education and Care Data Country Note: Chile.” OCED.

———. 2022. “Net Childcare Costs in EU Countries, 2021.” OCED. June 2022. https://www.oecd.org/els/soc/benefits-and-wages/Net%20childcare%20costs%20in%20EU%20countries_2021.pdf.

Paquette, Danielle. 2016. “Half of America’s Childcare Workers Need Food Stamps, Welfare Payments or Medicaid - The Washington Post.” July 11, 2016. https://www.washingtonpost.com/news/wonk/wp/2016/07/11/the-people-taking-care-of-our-kids-live-in-poverty/.

Québec, Institut de la statistique du. 2018. “Families by family structure and number of children aged of 24 and under, Québec, 1951-2016.” Institut de la Statistique du Québec. September 12, 2018. https://statistique.quebec.ca/en/document/families-in-quebec/tableau/families-by-family-structure-and-number-of-children-aged-of-24-and-under-quebec-1951-2016.

———. 2022a. “Enquête québécoise sur l’accessibilité et l’utilisation des services de garde 2021 - Faits saillants.” Institut de la Statistique du Québec. 2022. https://statistique.quebec.ca/fr/document/enquete-quebecoise-accessibilite-utilisation-services-garde-2021/publication/eqausg-apercu.

———. 2022b. “Accessibility and Use of Childcare Services in Québec in 2021.” Institut de La Statistique Du Québec. December 6, 2022. https://statistique.quebec.ca/en/communique/accessibility-use-childcare-services-quebec-2021.

Ruggles, Steven, Flood, Sarah, Sobek, Matthew, Brockman, Danika, Cooper, Grace, Richards, Stephanie, and Schouwiler, Megan. 2023. “IPUMS USA: Version 13.0.” Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V13.0.

SPLASH-db.eu. 2014. “Policy: ‘Family Policies: Italy’ (Information Provided by Guilia Ferrari).” SPLASH-Db.Eu. 2014. https://splash-db.eu/policydescription/family-policies-italy-2014/.

Statistics Canada. 2023. “Table 14-10-0287-03 Labour Force Characteristics by Province, Monthly, Seasonally Adjusted [Data Table].” April 6, 2023.

https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410028703.

Townsend, Peter, and David Gordon. 2022. *World Poverty: New Policies to Defeat an Old Enemy*. Bristol: Policy Press.

U.S. Bureau of Labor Statistics. 2022. “Women in the Labor Force: A Databook.” BLS Reports. March 2022. https://www.bls.gov/opub/reports/womens-databook/2021/home.htm.

Workman, Simon, and Steven Jessen-Howard. 2019. “New York State Cost of Quality Child Care Study.” Center for American Progress. https://raisingnewyork.org/wp-content/uploads/sites/2/2019/12/NY-Cost-of-Quality-Report-Raising-NY.pdf.

Wrohlich, Katharina. 2005. “The Excess Demand for Subsidized Child Care in Germany.” IZA Discussion paper series, No. 1515, Institute for the Study of Labor (IZA), Bonn. http://hdl.handle.net/10419/20814.

Zacharias, Ajit, Thomas Masterson, and Kijong Kim. 2009. “Distributional Impact of the American Recovery and Reinvestment Act: A Microsimulation Approach.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1420776.

APPENDIX A. DESCRIPTION OF THE SIMULATION

The likelihood of each person taking each occupation is computed

```
mlogit occ i.sex i.ageGroup i.educGroup i.raceGroup i.marstGroup i.childGroup
      i.nyc if donor [fw=intWeight], robust
```

Descriptions of the variables not found in ACS can be found in Appendix B. Predictions from this are put in *occLikelihood1…3* variables.

The likelihood of each person working in the child care sector at all is computed

```
probit hasCCJob i.ageGroup i.educGroup i.raceGroup i.marstGroup i.childGroup
      i.nyc if eligible & sex == 1 [pw=weight], robust
```

For each value of *sex*, 1 and 2. The predicted likelihoods are obtained from the model. A normal distribution fuzz is added to the predictions. The fuzz variance used was 0.2.

```
predict empLikelihoodMaleXb, xb
gen empLikelihoodMaleXbFuzzed = empLikelihoodMaleXb + rnormal(0, fuzzVariance)
gen empLikelihoodMale = normal(empLikelihoodMaleXbFuzzed)
```

Each sex receives the likelihoods from the model corresponding to their sex.

For each person and occupation, these likelihood values are multiplied together.

```
gen likelihoodEmpInOcc1 = empLikelihood * occLikelihood1
```

This is taken as a joint probability distribution representing the overall probability that a person will be selected for and take a job in that occupation. Eventually, for each occupation with jobs available, we will assign those jobs using random draws weighted by that joint probability. But first we calculate the inverse Mills Ratio and impute hours and earnings.

Inverse Mills Ratio (IMR) is calculated within sex and NYC/non-NYC cells. Within each cell, a probit model is run for child care labor force participation:

```
probit hasCCJob nchlt5 i.educGroup i.labforce i.ageGroup [fw=intWeight]
```

From this model the IMR is calculated within the cell

```
predict p if e(sample), xb
gen imr = normalden(p)/(normal(p))
```

At this point we run regressions to predict log-wages and weekly hours, again in sex and NYC/non-NYC cells. The wages regression is:

```
regress logWage age age2 i.educGroup i.childGroup occLikelihood* imr
      [aw=weight] if donor
```

The hours regression is:

```
regress hoursWorked age age2 i.educGroup i.childGroup i.labforce occLikelihood*
      predictedLogWage imr [aw=weight]
```

At this point is the unconstrained statistical matching using multiple imputation with hot-decking, and job allocation. This is done one occupation at a time. People in the recipient population are matched with similar people who have child care jobs. Matching is done in NYC/non-NYC cells. An implementation of multiple imputation with hot-decking, `mihd`, written by Thomas Masterson is used. Below, variables ending in *_sp* hold the values for the spouse, and are normalized as described in Appendix B. Variables which are used for matching are specified with *keepcat* or *keepcont*. Variables to be transferred from donors to recipients are specified in *cont* and *cat*. The *_sim* variables are initialized to be empty for all recipients and equal to the reported values for all donors.

```
mihd using jobs [aw=weight], id(year serial pernum)
      keepcat(labforce_sp educ_sp occLikelihood`o')
      keepcont(empLikelihood predictedLogWage predictedHours
            age age2 childGroup)
      cont(hourlyWage_sim hoursWorked_sim )
      cat(occ_sim)
      rep(1) imp(`variableWeights')
```

Variable weights are specified in Appendix B. The wage and hours imputed by the match may be interpreted as a hypothetical "job offer" made to each recipient. For each recipient we simulate whether or not they would be likely to accept the offer based on the hourly wage they reported for the prior year. We assume that people are unlikely to accept a job with a wage less than 75 percent of what they reported last year; some fuzz is applied to improve dispersion.

```
gen wouldTakeJob = (money_sim*wageIncrease > (rnormal(0.75, 0.15)*hourlyWage))
```

From among those who would take the job, we select those who receive jobs by simulating probability-weighted random draws. For the purpose of this simulation, `rdiscrete` was used, as described in the methodology section. Those attempting to reproduce this work should simply use `samplepps` instead, with *size* equal to *likelihoodEmpInOccN* (where N is whichever occupation is being allocated). When a recipient is drawn for job allocation, a number of jobs equal to that recipient's frequency weight are used up. Recipients are drawn until all new jobs are allocated. NYC and non-NYC jobs are in separate pools.

At the end, the wages of all people holding child care jobs are increased by the wage increase factor.

APPENDIX B. PARAMETERIZATION

Hourly earnings are the main wage variable used. These are computed as:

$$hourlyEarnings = incearn\,/\,(hoursWorked \times weeksWorked)$$

Where *incearn* is earned income. Earned incomes below $5000 and above $300,000 are dropped. Weeks worked is the exact number of weeks worked in the previous year, *wkswork1*, when available. Otherwise it is imputed from *wkswork2*, which gives the number of weeks as an interval, according to the mode values of *wkswork1* within each interval.

```
recode wkswork2 0=. 1=6 2=21 3=33 4=42 5=48 6=52, gen(wkswork2Recode)
gen weeksWorked = cond(mi(wkswork1), wkswork2Recode, wkswork1)
```

The eligible population is all unemployed people between the ages of 18 and 74 who reported either looking for work, being available for work, or being in the labor force. The cutoff age of 74 is determined heuristically from the distribution of ages of people who presently work in the child care sector.

The jobs considered are in industry 8470, "Child Care Services," occupations 230 ("Education and childcare administrators"), 2545 ("Teaching assistants"), and 4600 ("Childcare workers"). Whether or not a person is considered in New York City is according to when the *city* variable takes the value 4610.

Marital status is grouped from *marst* to take values "Married," "Separated/Widowed/Divorced," and "Single." Number of children under 5 is grouped from *nchlt5* to the values 0, 1, and 2+. Ages are grouped into those under 25, those 26-40, those 41-65, and those 66 and over.

Variable weights for multiple imputation with hot-decking via `mihd` are:

labforce_sp	4
educ_sp	4
occLikelihood`o'	1000
empLikelihood	100
imputedLogWage	500
imputedHours	500
age	20
age2	20
childGroup	20

APPENDIX C. SIMULATION RESULTS

Donors: 63,465. Recipients: 769,725.

Donor occupations: 5,141 administrators, 11,583 teaching assistants, 46,741 child care workers. Simulated occupations: 22,036 administrators, 52,434 teaching assistants, 195,168 child care workers (for scenarios with 75 percent uptake).

```
mlogit occ i.sex i.ageGroup i.educGroup i.raceGroup i.marstGroup i.childGroup
      i.nyc if donor [fw=intWeight], robust
```

Multinomial logistic regression		Number of obs	63,465			
		Wald chi2(28)	8541.37			

		Prob > chi2	0			
Log pseudolikelihood	-42396.109	Pseudo R2	0.0964			
occ 1						
	Coefficient	Robust std. err.	z	P>z [95%	conf.	interval]
sex						
female	-0.3703446	0.0616162	-6.01	0	-0.4911102	-0.2495791
ageGroup						
26-40	1.803638	0.0776217	23.24	0	1.651502	1.955774
41-65	1.598165	0.0807428	19.79	0	1.439912	1.756418
66+	1.997937	0.0991985	20.14	0	1.803511	2.192362
educGroup						
Completed HS/GED	0.7179983	0.0921588	7.79	0	0.5373703	0.8986263
Completed bachelor's	2.642636	0.0908396	29.09	0	2.464593	2.820678
raceGroup						
Black not hispanic	-0.6485346	0.0484196	-13.39	0	-0.7434353	-0.5536339
Other not hispanic	-0.2568006	0.0572713	-4.48	0	-0.3690502	-0.144551
Hispanic	-0.9086357	0.0448656	-20.25	0	-0.9965707	-0.8207006
marstGroup						

Separated/Widowed/Divorced	-0.2832199	0.0458748	-6.17	0	-0.373133	-0.1933069
Single	-0.3715506	0.0423833	-8.77	0	-0.4546204	-0.2884808
childGroup						
1	-0.0913871	0.0664747	-1.37	0.169	-0.221675	0.0389009
2+	-0.1875606	0.0956818	-1.96	0.05	-0.3750934	-0.0000277
1.nyc	-0.0227831	0.0372713	-0.61	0.541	-0.0958336	0.0502674
_cons	-4.218178	0.1352698	-31.18	0	-4.483302	-3.953054
occ 2						
sex						
female	0.8245635	0.0588107	14.02	0	0.7092966	0.9398303
ageGroup						
26-40	0.1077564	0.032654	3.3	0.001	0.0437557	0.1717572
41-65	-0.3036239	0.0343649	-8.84	0	-0.3709778	-0.2362699
66+	-0.0326962	0.058346	-0.56	0.575	-0.1470523	0.0816599
educGroup						
Completed HS/GED	0.9516764	0.0506162	18.8	0	0.8524706	1.050882
Completed bachelor's	1.215569	0.0541802	22.44	0	1.109378	1.32176
raceGroup						
Black not hispanic	-0.45109	0.03349	-13.47	0	-0.5167293	-0.3854508

Other not hispanic	0.0094169	0.0421306	0.22	0.823	-0.0731575	0.0919914
Hispanic	-0.459799	0.0313601	-14.66	0	-0.5212637	-0.3983343
marstGroup						
Separated/Widowed/Divorced	-0.2056615	0.0325117	-6.33	0	-0.2693833	-0.1419396
Single	-0.2220727	0.025692	-8.64	0	-0.2724281	-0.1717173
childGroup						
1	0.4253176	0.0385362	11.04	0	0.3497879	0.5008472
2+	-0.5109644	0.0868275	-5.88	0	-0.6811431	-0.3407856
1.nyc	-0.3180184	0.0270796	-11.74	0	-0.3710935	-0.2649433
_cons	-2.546996	0.0819636	-31.07	0	-2.707641	-2.38635
3 | (base outcome)						

```
probit hasCCJob i.ageGroup i.educGroup i.raceGroup i.marstGroup i.childGroup
      i.nyc if eligible & sex == 1 [pw=weight], robust
```

Probit regression			Number of obs	18,171		
			Wald chi2(13)	31.63		
			Prob > chi2	0.0027		
Log pseudolikelihood	-21419.485		Pseudo R2	0.0189		
		Robust				
hasCCJob	Coefficient	std. err.	z	P>z [95%	conf.	interval]
ageGroup						

26-40	0.0485401	0.0997932	0.49	0.627	-0.147051	0.2441311
41-65	-0.1515922	0.1015189	-1.49	0.135	-0.3505656	0.0473811
66+	-0.4793076	0.1873607	-2.56	0.011	-0.8465279	-0.1120874
educGroup						
Completed HS/GED	0.3381875	0.1847994	1.83	0.067	-0.0240126	0.7003876
Completed bachelor's	0.2616447	0.1911038	1.37	0.171	-0.1129117	0.6362012
raceGroup						
Black not hispanic	0.0862047	0.0995262	0.87	0.386	-0.108863	0.2812725
Other not hispanic	0.0593934	0.1206991	0.49	0.623	-0.1771725	0.2959593
Hispanic	0.0896557	0.0996866	0.9	0.368	-0.1057265	0.285038
marstGroup						
Separated/Widowed/Divorced	-0.2613507	0.1684899	-1.55	0.121	-0.5915849	0.0688835
Single	-0.1445649	0.0943525	-1.53	0.125	-0.3294924	0.0403625
childGroup						
1	0.1464919	0.1889973	0.78	0.438	-0.2239361	0.5169198
2+	-0.100236	0.2335142	-0.43	0.668	-0.5579155	0.3574435
1.nyc	0.0281696	0.0776189	0.36	0.717	-0.1239606	0.1802997
_cons	-2.552338	0.216727	-11.78	0	-2.977115	-2.127561

probit hasCCJob nchlt5 i.educGroup i.labforce i.ageGroup [fw=intWeight]

```
Labor force probit: nyc = 0, sex = 1

Probit regression                              Number of obs = 198,832
                                               LR chi2(7)     =   612.73
                                               Prob > chi2    =   0.0000
```

Log likelihood = -9557.4385 Pseudo R2 = 0.0311

hasCCJob	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
nchlt5	-.0203451	.029687	-0.69	0.493	-.0785306	.0378403
d_educGroup_1	.5092154	.0575667	8.85	0.000	.3963867	.6220441
d_educGroup_2	.6334932	.0589876	10.74	0.000	.5178796	.7491068
d_labforce__1	-.0949677	.0249465	-3.81	0.000	-.143862	-.0460734
d_ageGroup_1	.1323295	.0218184	6.07	0.000	.0895662	.1750928
d_ageGroup_2	-.2885946	.0266422	-10.83	0.000	-.3408124	-.2363768
d_ageGroup_3	-.3840707	.0453931	-8.46	0.000	-.4730396	-.2951018
_cons	-2.823769	.0577733	-48.88	0.000	-2.937003	-2.710536

Labor force probit: nyc = 0, sex = 2

Probit regression Number of obs = 206,740
LR chi2(7) = 2558.26
Prob > chi2 = 0.0000
Log likelihood = -81955.807 Pseudo R2 = 0.0154

hasCCJob	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
nchlt5	-.1310913	.0089853	-14.59	0.000	-.1487021	-.1134805
d_educGroup_1	.0818799	.0143925	5.69	0.000	.0536711	.1100887
d_educGroup_2	-.1901216	.0155349	-12.24	0.000	-.2205695	-.1596737
d_labforce__1	.2031412	.0076959	26.40	0.000	.1880575	.2182249
d_ageGroup_1	.040294	.0100779	4.00	0.000	.0205417	.0600463
d_ageGroup_2	.0453568	.009336	4.86	0.000	.0270586	.0636551
d_ageGroup_3	-.2722524	.0158216	-17.21	0.000	-.3032622	-.2412426
_cons	-1.164017	.015031	-77.44	0.000	-1.193477	-1.134556

Labor force probit: nyc = 1, sex = 1

Probit regression Number of obs = 210,125
LR chi2(7) = 439.87
Prob > chi2 = 0.0000
Log likelihood = -11807.392 Pseudo R2 = 0.0183

hasCCJob	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
nchlt5	.1169902	.0203862	5.74	0.000	.077034	.1569464

```
d_educGroup_1 |   .2863053   .0300956      9.51   0.000  .2273191      .3452916
d_educGroup_2 |   .0146814   .0339035      0.43   0.665  -.0517683      .081131
d_labforce__1 |   .1675778   .0201341      8.32   0.000  .1281157      .2070399
 d_ageGroup_1 |   .0374761   .022216       1.69   0.092  -.0060665      .0810188
 d_ageGroup_2 |  -.0117923   .0231182      -0.51   0.610 -.0571031      .0335184
 d_ageGroup_3 |  -.4333811   .0743313      -5.83   0.000 -.5790678   -.2876944
       _cons |  -2.567926   .0327628   -78.38   0.000     -2.63214   -2.503712
------------------------------------------------------------------------------
(29526 missing values generated)
(7,974 real changes made)
Labor force probit: nyc = 1, sex = 2

Probit regression                                       Number of obs = 217,493
                                                        LR chi2(7)     = 7056.14
                                                        Prob > chi2    =  0.0000
Log likelihood = -85390.412                             Pseudo R2      =  0.0397

------------------------------------------------------------------------------
       hasCCJob | Coefficient  Std. err.    z      P>|z|  [95% conf. interval]
--------------+---------------------------------------------------------------
       nchlt5 |  -.0738365   .0095369      -7.74   0.000 -.0925285   -.0551446
d_educGroup_1 |  -.0054398   .0102703      -0.53   0.596 -.0255692      .0146895
d_educGroup_2 |  -.3130516   .0114644    -27.31   0.000    -.3355215   -.2905817
d_labforce__1 |      .125228   .0076326    16.41   0.000 .1102684      .1401877
 d_ageGroup_1 |   .4740669   .0112212      42.25   0.000 .4520737      .4960601
 d_ageGroup_2 |   .6956142   .0108896      63.88   0.000 .6742709      .7169575
 d_ageGroup_3 |   .6390765   .0186279      34.31   0.000 .6025665      .6755865
       _cons |   -1.51005   .0131211  -115.09   0.000     -1.535767   -1.484333
------------------------------------------------------------------------------
```

regress logWage age age2 i.educGroup i.childGroup occLikelihood* imr [aw=weight] if donor

regress hoursWorked age age2 i.educGroup i.childGroup i.labforce occLikelihood* predictedLogWage imr [aw=weight]

```
Wage and hours imputation: nyc = 0, sex = 1

(sum of wgt is 1,716)
note: occLikelihood2 omitted because of collinearity.

       Source |       SS            df      MS      Number of obs   =     79
-------------+----------------------------------   F(9, 69)      =      0.76
       Model |  2.31492423          9  .257213803   Prob > F      =      0.6536
       Residual |  23.3690225       69  .338681485   R-squared    =      0.0901
```

```
-------------+----------------------------------   Adj R-squared   =   -0.0285
       Total |  25.6839467          78  .329281368   Root MSE    =      .58196

--------------------------------------------------------------------------------
       logWage | Coefficient  Std. err.    t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |   .0001899      .03152  0.01   0.995   -.0626908      .0630706
       age2 |   .0000317   .0003915 0.08   0.936   -.0007492      .0008127
 d_educGroup_1 |   .9036283   .6306152      1.43   0.156   -.3544146      2.161671
 d_educGroup_2 |   .9776623   .7212917      1.36   0.180   -.4612752      2.4166
 d_childGrou_1 |   .3100989   .3338133      0.93   0.356   -.3558406      .9760383
 d_childGrou_2 |   .6515597   .8111213      0.80   0.425   -.966583       2.269702
occLikelihood1 |   5.594642  3.26036       1.72   0.091   -.9095974      12.09888
occLikelihood2 |            0  (omitted)
occLikelihood3 |   6.457084   3.295332      1.96   0.054   -.1169234      13.03109
       lambda |  -.2533635   .9100967      -0.28   0.782  -2.068957      1.56223
       _cons |  -3.411974   3.584757-0.95   0.345  -10.56337      3.739419
--------------------------------------------------------------------------------
(sum of wgt is 1,716)

       Source |       SS           df      MS      Number of obs   =     79
-------------+----------------------------------   F(9, 69)       =      0.76
       Model |  2.31492423          9  .257213803   Prob > F      =      0.6536
       Residual |  23.3690225       69  .338681485   R-squared    =      0.0901
-------------+----------------------------------   Adj R-squared   =   -0.0285
       Total |  25.6839467          78  .329281368   Root MSE    =      .58196

--------------------------------------------------------------------------------
       logWage | Coefficient  Std. err.    t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |   .0001899      .03152  0.01   0.995   -.0626908      .0630706
       age2 |   .0000317   .0003915 0.08   0.936   -.0007492      .0008127
 d_educGroup_1 |   .9036283   .6306152      1.43   0.156   -.3544146      2.161671
 d_educGroup_2 |   .9776623   .7212917      1.36   0.180   -.4612752      2.4166
 d_childGrou_1 |   .3100989   .3338133      0.93   0.356   -.3558406      .9760383
 d_childGrou_2 |   .6515597   .8111213      0.80   0.425   -.966583       2.269702
occLikelihood1 |   5.594642  3.26036       1.72   0.091   -.9095974      12.09888
occLikelihood3 |   6.457084   3.295332      1.96   0.054   -.1169234      13.03109
       lambda |  -.2533635   .9100967      -0.28   0.782  -2.068957      1.56223
       _cons |  -3.411974   3.584757-0.95   0.345  -10.56337      3.739419
--------------------------------------------------------------------------------
(27,302 missing values generated)
(10,198 real changes made)
(sum of wgt is 1,716)
note: occLikelihood2 omitted because of collinearity.
note: occLikelihood3 omitted because of collinearity.

       Source |       SS           df      MS      Number of obs   =     79
```

```
-------------+----------------------------------   F(10, 68)      =      1.79
       Model |  2760.98362          10  276.098362   Prob > F       =    0.0788
    Residual |  10479.5348          68  154.110806   R-squared      =    0.2085
-------------+----------------------------------   Adj R-squared   =    0.0921
       Total |  13240.5184          78  169.750236   Root MSE       =    12.414

--------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t    P>|t|   [95% conf. interval]
---------------+----------------------------------------------------------------
           age |   .5544835   .6852183   0.81   0.421   -.8128484    1.921815
          age2 |  -.0084974   .0085677  -0.99   0.325   -.025594     .0085992
 d_educGroup_1 |   10.90849   15.72923   0.69   0.490   -20.47869    42.29567
 d_educGroup_2 |   18.71717   17.96697   1.04   0.301   -17.13536    54.5697
 d_childGrou_1 |   9.102962   6.428828   1.42   0.161   -3.725563    21.93149
 d_childGrou_2 |  -5.859712   20.49762  -0.29   0.776   -46.76208    35.04266
 d_labforce__1 |   2.097348   4.527543   0.46   0.645   -6.937223    11.13192
occLikelihood1 |  -4.101894   25.37887  -0.16   0.872   -54.74463    46.54085
occLikelihood2 |          0  (omitted)
occLikelihood3 |          0  (omitted)
imputedLogWage |   20.70232   11.06466   1.87   0.066   -1.376872    42.78151
        lambda |   30.10933   19.49346   1.54   0.127   -8.789255    69.00792
         _cons |  -118.6615   61.06324  -1.94   0.056   -240.5112    3.18834
--------------------------------------------------------------------------------
(sum of wgt is 1,716)

      Source |       SS           df       MS      Number of obs   =        79
-------------+----------------------------------   F(10, 68)      =      1.79
       Model |  2760.98362          10  276.098362   Prob > F       =    0.0788
    Residual |  10479.5348          68  154.110806   R-squared      =    0.2085
-------------+----------------------------------   Adj R-squared   =    0.0921
       Total |  13240.5184          78  169.750236   Root MSE       =    12.414

--------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t    P>|t|   [95% conf. interval]
---------------+----------------------------------------------------------------
           age |   .5544835   .6852183   0.81   0.421   -.8128484    1.921815
          age2 |  -.0084974   .0085677  -0.99   0.325   -.025594     .0085992
 d_educGroup_1 |   10.90849   15.72923   0.69   0.490   -20.47869    42.29567
 d_educGroup_2 |   18.71717   17.96697   1.04   0.301   -17.13536    54.5697
 d_childGrou_1 |   9.102962   6.428828   1.42   0.161   -3.725563    21.93149
 d_childGrou_2 |  -5.859712   20.49762  -0.29   0.776   -46.76208    35.04266
 d_labforce__1 |   2.097348   4.527543   0.46   0.645   -6.937223    11.13192
occLikelihood1 |  -4.101894   25.37887  -0.16   0.872   -54.74463    46.54085
imputedLogWage |   20.70232   11.06466   1.87   0.066   -1.376872    42.78151
        lambda |   30.10933   19.49346   1.54   0.127   -8.789255    69.00792
         _cons |  -118.6615   61.06324  -1.94   0.056   -240.5112    3.18834
--------------------------------------------------------------------------------
(27,302 missing values generated)
```

```
(10,198 real changes made)
Wage and hours imputation: nyc = 0, sex = 2

(sum of wgt is 28,676)
note: occLikelihood2 omitted because of collinearity.

       Source |       SS            df       MS      Number of obs   =     1,357
-------------+----------------------------------   F(9, 1347)      =     12.63
       Model |  42.3107736          9  4.70119706   Prob > F        =    0.0000
    Residual |  501.190815      1,347  .372079299   R-squared       =    0.0778
-------------+----------------------------------   Adj R-squared   =    0.0717
       Total |  543.501589     1,356  .400812381   Root MSE        =    .60998

------------------------------------------------------------------------------
       logWage | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
---------------+----------------------------------------------------------------
           age |   -.024549   .0090662    -2.71   0.007    -.0423343   -.0067637
          age2 |   .0002691   .0001014     2.65   0.008     .0000701    .0004681
 d_educGroup_1 |   .2345374   .0909872     2.58   0.010     .0560453    .4130295
 d_educGroup_2 |   .3859731   .1268115     3.04   0.002     .1372036    .6347426
 d_childGrou_1 |   .3843477   .0835107     4.60   0.000     .2205225     .548173
 d_childGrou_2 |  -.0093965   .1309864    -0.07   0.943     -.266356    .2475631
occLikelihood1 |   3.270609   .6121341     5.34   0.000     2.069769    4.471449
occLikelihood2 |          0  (omitted)
occLikelihood3 |   1.910293   .4388653     4.35   0.000     1.049359    2.771226
        lambda |    -.25344   .2116328    -1.20   0.231    -.6686057    .1617257
         _cons |    1.49431   .5115245     2.92   0.004      .490839    2.497781
------------------------------------------------------------------------------
(sum of wgt is 28,676)

       Source |       SS            df       MS      Number of obs   =     1,357
-------------+----------------------------------   F(9, 1347)      =     12.63
       Model |  42.3107736          9  4.70119706   Prob > F        =    0.0000
    Residual |  501.190815      1,347  .372079299   R-squared       =    0.0778
-------------+----------------------------------   Adj R-squared   =    0.0717
       Total |  543.501589     1,356  .400812381   Root MSE        =    .60998

------------------------------------------------------------------------------
       logWage | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
---------------+----------------------------------------------------------------
           age |   -.024549   .0090662    -2.71   0.007    -.0423343   -.0067637
          age2 |   .0002691   .0001014     2.65   0.008     .0000701    .0004681
 d_educGroup_1 |   .2345374   .0909872     2.58   0.010     .0560453    .4130295
 d_educGroup_2 |   .3859731   .1268115     3.04   0.002     .1372036    .6347426
 d_childGrou_1 |   .3843477   .0835107     4.60   0.000     .2205225     .548173
 d_childGrou_2 |  -.0093965   .1309864    -0.07   0.943     -.266356    .2475631
occLikelihood1 |   3.270609   .6121341     5.34   0.000     2.069769    4.471449
occLikelihood3 |   1.910293   .4388653     4.35   0.000     1.049359    2.771226
```

```
      lambda |      -.25344   .2116328    -1.20   0.231  -.6686057      .1617257
       _cons |1.49431   .5115245    2.92   0.004   .4908392.497781
--------------------------------------------------------------------------------
(26,907 missing values generated)
(10,593 real changes made)
(sum of wgt is 28,676)
note: occLikelihood1 omitted because of collinearity.
note: occLikelihood2 omitted because of collinearity.

      Source |       SS            df       MS      Number of obs   =     1,357
-------------+----------------------------------   F(10, 1346)    =     11.11
       Model |  17355.9379          10  1735.59379   Prob > F    =      0.0000
    Residual |  210346.672       1,346  156.275388   R-squared       =      0.0762
-------------+----------------------------------   Adj R-squared   =     0.0694
       Total |   227702.61   1,356  167.922279   Root MSE        =      12.501

--------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
         age |   .9062619   .1951331  4.64   0.000   .5234638      1.28906
        age2 |  -.0085356   .0023609 -3.62   0.000  -.013167   -.0039043
 d_educGroup_1 |   6.338323   1.838698      3.45   0.001   2.731298      9.945348
 d_educGroup_2 |   12.21613   3.210349      3.81   0.000   5.918298      18.51396
 d_childGrou_1 |   4.832005   1.924657      2.51   0.012   1.056351      8.607659
 d_childGrou_2 |  -2.237161   3.417607      -0.65   0.513  -8.941577      4.467255
 d_labforce__1 |  -1.441872   1.741374      -0.83   0.408  -4.857974      1.97423
occLikelihood1 |            0  (omitted)
occLikelihood2 |            0  (omitted)
occLikelihood3 |   20.64677   5.587389      3.70   0.000   9.685832      31.60771
imputedLogWage |   -3.90424   3.870119      -1.01   0.313  -11.49636      3.68788
      lambda |  -8.347815   9.53455-0.88   0.381  -27.05201      10.35638
       _cons |   17.86973   22.89959       0.78   0.435  -27.05304      62.79249
--------------------------------------------------------------------------------
(sum of wgt is 28,676)

      Source |       SS            df       MS      Number of obs   =     1,357
-------------+----------------------------------   F(10, 1346)    =     11.11
       Model |  17355.9379          10  1735.59379   Prob > F    =      0.0000
    Residual |  210346.672       1,346  156.275388   R-squared       =      0.0762
-------------+----------------------------------   Adj R-squared   =     0.0694
       Total |   227702.61   1,356  167.922279   Root MSE        =      12.501

--------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
         age |   .9062619   .1951331  4.64   0.000   .5234638      1.28906
        age2 |  -.0085356   .0023609 -3.62   0.000  -.013167   -.0039043
 d_educGroup_1 |   6.338323   1.838698      3.45   0.001   2.731298      9.945348
```

```
 d_educGroup_2 |   12.21613   3.210349      3.81   0.000   5.918298      18.51396
 d_childGrou_1 |   4.832005   1.924657      2.51   0.012   1.056351      8.607659
 d_childGrou_2 |  -2.237161   3.417607      -0.65   0.513  -8.941577     4.467255
 d_labforce__1 |  -1.441872   1.741374      -0.83   0.408  -4.857974     1.97423
occLikelihood3 |   20.64677   5.587389      3.70   0.000   9.685832      31.60771
imputedLogWage |   -3.90424   3.870119      -1.01   0.313  -11.49636     3.68788
       lambda |  -8.347815   9.53455-0.88   0.381  -27.05201     10.35638
        _cons |   17.86973   22.89959      0.78   0.435   -27.05304     62.79249
-------------------------------------------------------------------------------
(26,907 missing values generated)
(10,593 real changes made)
Wage and hours imputation: nyc = 1, sex = 1

(sum of wgt is 2,158)
note: occLikelihood2 omitted because of collinearity.

       Source |       SS            df       MS      Number of obs   =      77
-------------+----------------------------------   F(9, 67)        =      5.41
       Model |   15.764345          9  1.75159389   Prob > F        =    0.0000
       Residual |  21.6915391       67  .323754315   R-squared       =    0.4209
-------------+----------------------------------   Adj R-squared   =    0.3431
       Total |  37.4558842          76  .492840581   Root MSE        =    .56899

-------------------------------------------------------------------------------
       logWage | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+---------------------------------------------------------------
       age |   .0878637   .0435973  2.02   0.048   .0008431      .1748843
       age2 |   -.000954     .000513-1.86   0.067  -.001978       .00007
 d_educGroup_1 |   .9251704   .4182201      2.21   0.030   .0903995      1.759941
 d_educGroup_2 |   1.024323   .6212272      1.65   0.104   -.2156514     2.264298
 d_childGrou_1 |   .5694315   .3042209      1.87   0.066   -.037796      1.176659
 d_childGrou_2 |   .8195546   .6252045      1.31   0.194   -.4283588     2.067468
occLikelihood1 |  -3.733751   4.304407      -0.87   0.389  -12.32538     4.857882
occLikelihood2 |            0  (omitted)
occLikelihood3 |  -4.330934   3.935194      -1.10   0.275  -12.18561     3.523747
       lambda |   .5572738   .9480857      0.59   0.559   -1.335113     2.449661
        _cons |   2.411059   4.513622      0.53   0.595   -6.598168     11.42029
-------------------------------------------------------------------------------
(sum of wgt is 2,158)

       Source |       SS            df       MS      Number of obs   =      77
-------------+----------------------------------   F(9, 67)        =      5.41
       Model |   15.764345          9  1.75159389   Prob > F        =    0.0000
       Residual |  21.6915391       67  .323754315   R-squared       =    0.4209
-------------+----------------------------------   Adj R-squared   =    0.3431
       Total |  37.4558842          76  .492840581   Root MSE        =    .56899

-------------------------------------------------------------------------------
```

```
       logWage | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |   .0878637   .0435973  2.02   0.048   .0008431      .1748843
       age2 |   -.000954     .000513-1.86   0.067  -.001978      .00007
 d_educGroup_1 |   .9251704   .4182201      2.21   0.030  .0903995      1.759941
 d_educGroup_2 |   1.024323   .6212272      1.65   0.104  -.2156514      2.264298
 d_childGrou_1 |   .5694315   .3042209      1.87   0.066  -.037796      1.176659
 d_childGrou_2 |   .8195546   .6252045      1.31   0.194  -.4283588      2.067468
occLikelihood1 |  -3.733751   4.304407      -0.87   0.389  -12.32538      4.857882
occLikelihood3 |  -4.330934   3.935194      -1.10   0.275  -12.18561      3.523747
       lambda |   .5572738   .9480857      0.59   0.559  -1.335113      2.449661
       _cons |   2.411059   4.513622       0.53   0.595  -6.598168      11.42029
--------------------------------------------------------------------------------
(29,526 missing values generated)
(7,974 real changes made)
(sum of wgt is 2,158)
note: occLikelihood2 omitted because of collinearity.
note: occLikelihood3 omitted because of collinearity.

       Source |       SS            df      MS     Number of obs   =     77
-------------+----------------------------------  F(10, 66)      =      1.99
       Model |  1567.52355          10  156.752355   Prob > F      =      0.0486
       Residual |  5202.71331       66  78.8289895   R-squared     =      0.2315
-------------+----------------------------------  Adj R-squared   =     0.1151
       Total |  6770.23686          76  89.0820639   Root MSE      =      8.8786

--------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |   .2972312   1.394752  0.21   0.832  -2.48748      3.081943
       age2 |  -.0025204   .0159992 -0.16   0.875  -.0344638      .0294231
 d_educGroup_1 |  -8.207841   17.18435      -0.48   0.634  -42.51751      26.10183
 d_educGroup_2 |  -8.455696   19.12572      -0.44   0.660  -46.64142      29.73003
 d_childGrou_1 |  -4.538597   10.61764      -0.43   0.670  -25.7374      16.66021
 d_childGrou_2 |      -9.4167   14.58051      -0.65   0.521  -38.52762      19.69422
 d_labforce__1 |      .614146   5.265188      0.12   0.907  -9.898141      11.12643
occLikelihood1 |   8.275296   25.2446       0.33   0.744  -42.12717      58.67776
occLikelihood2 |            0  (omitted)
occLikelihood3 |            0  (omitted)
imputedLogWage |   10.08334   14.19524      0.71   0.480  -18.25836      38.42505
       lambda |  -4.355826   30.12149      -0.14   0.885  -64.49532      55.78367
       _cons |   20.45926   90.85104       0.23   0.823  -160.9307      201.8492
--------------------------------------------------------------------------------
(sum of wgt is 2,158)

       Source |       SS            df      MS     Number of obs   =     77
-------------+----------------------------------  F(10, 66)      =      1.99
       Model |  1567.52355          10  156.752355   Prob > F      =      0.0486
```

```
       Residual |  5202.71331        66  78.8289895   R-squared   =     0.2315
-------------+----------------------------------   Adj R-squared   =     0.1151
       Total |  6770.23686          76  89.0820639   Root MSE    =     8.8786

--------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |   .2972312   1.394752  0.21   0.832   -2.48748      3.081943
       age2 |  -.0025204   .0159992 -0.16   0.875  -.0344638     .0294231
 d_educGroup_1 |  -8.207841   17.18435      -0.48   0.634  -42.51751     26.10183
 d_educGroup_2 |  -8.455696   19.12572      -0.44   0.660  -46.64142     29.73003
 d_childGrou_1 |  -4.538597   10.61764      -0.43   0.670  -25.7374      16.66021
 d_childGrou_2 |      -9.4167   14.58051    -0.65   0.521  -38.52762     19.69422
 d_labforce__1 |      .614146   5.265188    0.12   0.907   -9.898141     11.12643
occLikelihood1 |   8.275296  25.2446       0.33   0.744   -42.12717     58.67776
imputedLogWage |   10.08334   14.19524      0.71   0.480   -18.25836     38.42505
       lambda |  -4.355826   30.12149      -0.14   0.885  -64.49532     55.78367
       _cons |   20.45926   90.85104       0.23   0.823   -160.9307     201.8492
--------------------------------------------------------------------------------
(29,526 missing values generated)
(7,974 real changes made)
Wage and hours imputation: nyc = 1, sex = 2

(sum of wgt is 30,915)
note: occLikelihood2 omitted because of collinearity.

       Source |       SS            df      MS     Number of obs   =     1,098
-------------+----------------------------------   F(9, 1088)      =     13.14
       Model |  46.2985735          9  5.14428594   Prob > F       =     0.0000
       Residual |   425.94521       1,088  .391493759   R-squared       =      0.0980
-------------+----------------------------------   Adj R-squared   =     0.0906
       Total |  472.243783   1,097  .430486585   Root MSE        =     .62569

--------------------------------------------------------------------------------
       logWage | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |  -.0427335   .0152142  -2.81   0.005  -.0725859     -.012881
       age2 |   .0004709     .000151       3.12   0.002   .0001746       .0007672
 d_educGroup_1 |   .1016838   .0783638      1.30   0.195   -.0520774      .255445
 d_educGroup_2 |   .3518862   .1526861      2.30   0.021   .0522937       .6514787
 d_childGrou_1 |      -.1593   .0917711     -1.74   0.083  -.3393685      .0207684
 d_childGrou_2 |  -.1969775  .158855-1.24   0.215  -.5086743     .1147193
occLikelihood1 |  -1.615096   1.038225      -1.56   0.120  -3.652245      .422053
occLikelihood2 |             0  (omitted)
occLikelihood3 |  -2.167686   .6914585      -3.13   0.002  -3.524429   -.8109433
       lambda |  -.4664746   .2409682      -1.94   0.053  -.9392896      .0063405
       _cons |   5.868822   .9078474       6.46   0.000   4.087493      7.650152
--------------------------------------------------------------------------------
```

```
(sum of wgt is 30,915)

       Source |       SS             df      MS      Number of obs   =     1,098
-------------+----------------------------------   F(9, 1088)     =      13.14
       Model |  46.2985735          9  5.14428594   Prob > F      =      0.0000
       Residual |   425.94521       1,088  .391493759   R-squared       =      0.0980
-------------+----------------------------------   Adj R-squared   =     0.0906
       Total |  472.243783   1,097  .430486585   Root MSE        =      .62569

--------------------------------------------------------------------------------
       logWage | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |  -.0427335   .0152142  -2.81   0.005  -.0725859      -.012881
       age2 |   .0004709     .000151       3.12   0.002   .0001746       .0007672
 d_educGroup_1 |   .1016838   .0783638      1.30   0.195  -.0520774      .255445
 d_educGroup_2 |   .3518862   .1526861      2.30   0.021   .0522937       .6514787
 d_childGrou_1 |      -.1593   .0917711     -1.74   0.083  -.3393685      .0207684
 d_childGrou_2 |  -.1969775  .158855-1.24   0.215  -.5086743      .1147193
occLikelihood1 |  -1.615096   1.038225      -1.56   0.120  -3.652245      .422053
occLikelihood3 |  -2.167686   .6914585      -3.13   0.002  -3.524429   -.8109433
       lambda |  -.4664746   .2409682      -1.94   0.053  -.9392896      .0063405
       _cons |   5.868822   .9078474       6.46   0.000   4.087493       7.650152
--------------------------------------------------------------------------------
(28,765 missing values generated)
(8,735 real changes made)
(sum of wgt is 30,915)
note: occLikelihood2 omitted because of collinearity.
note: occLikelihood3 omitted because of collinearity.

       Source |       SS             df      MS      Number of obs   =     1,098
-------------+----------------------------------   F(10, 1087)    =      4.25
       Model |  5718.49167          10  571.849167   Prob > F    =      0.0000
       Residual |  146244.088       1,087  134.539179   R-squared       =      0.0376
-------------+----------------------------------   Adj R-squared   =     0.0288
       Total |   151962.58   1,097  138.525597   Root MSE        =      11.599

--------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+----------------------------------------------------------------
       age |  -.7008636   .4071499  -1.72   0.085  -1.499752      .098025
       age2 |   .0077243   .0041937 1.84   0.066  -.0005042      .0159529
 d_educGroup_1 |   6.072999   2.001066      3.03   0.002   2.146611       9.999388
 d_educGroup_2 |   17.79503   3.718593      4.79   0.000   10.498625.09147
 d_childGrou_1 |   1.012298   1.426642      0.71   0.478  -1.786986       3.811582
 d_childGrou_2 |  -2.758281  3.54734-0.78   0.437  -9.718689      4.202128
 d_labforce__1 |  -.5187321   .9819707      -0.53   0.597  -2.445505      1.408041
occLikelihood1 |  -5.923235   12.14448      -0.49   0.626  -29.75252      17.90605
occLikelihood2 |            0  (omitted)
```

```
occLikelihood3 |             0  (omitted)
imputedLogWage |  -17.74009   6.302541      -2.81   0.005  -30.10662   -5.373571
       lambda |  -14.89377   5.468522      -2.72   0.007  -25.62382   -4.163712
       _cons |   111.7589   26.64748       4.19   0.000   59.47258      164.0452
------------------------------------------------------------------------------
(sum of wgt is 30,915)

       Source |       SS            df       MS      Number of obs   =     1,098
-------------+----------------------------------   F(10, 1087)    =      4.25
       Model |  5718.49167          10  571.849167   Prob > F    =      0.0000
       Residual |  146244.088       1,087  134.539179   R-squared       =      0.0376
-------------+----------------------------------   Adj R-squared   =     0.0288
       Total |   151962.58   1,097  138.525597   Root MSE        =      11.599

------------------------------------------------------------------------------
   hoursWorked | Coefficient  Std. err.     t      P>|t|  [95% conf. interval]
---------------+--------------------------------------------------------------
       age |  -.7008636   .4071499  -1.72   0.085  -1.499752      .098025
       age2 |   .0077243   .0041937 1.84   0.066   -.0005042      .0159529
 d_educGroup_1 |   6.072999   2.001066      3.03   0.002   2.146611      9.999388
 d_educGroup_2 |   17.79503   3.718593      4.79   0.000   10.4986 25.09147
 d_childGrou_1 |   1.012298   1.426642      0.71   0.478   -1.786986     3.811582
 d_childGrou_2 |  -2.758281  3.54734 -0.78   0.437  -9.718689     4.202128
 d_labforce__1 |  -.5187321   .9819707      -0.53   0.597  -2.445505     1.408041
occLikelihood1 |  -5.923235   12.14448      -0.49   0.626  -29.75252     17.90605
imputedLogWage |  -17.74009   6.302541      -2.81   0.005  -30.10662   -5.373571
       lambda |  -14.89377   5.468522      -2.72   0.007  -25.62382   -4.163712
       _cons |   111.7589   26.64748       4.19   0.000   59.47258      164.0452
------------------------------------------------------------------------------
```