A Welfare Analysis of Universal Childcare: Lessons From a Canadian Reform*

Sébastien Montpetit† Luisa Carrer‡ Pierre-Loup Beauregard§

June 13, 2025

Abstract

We assess the welfare impact of the introduction of universal daycare services in Québec in 1997. Unlike the standard sufficient-statistic metric, which assumes marginal changes in fiscal policy, our approach accounts for the non-marginal nature of the program and quantifies non-pecuniary benefits. Through a structural model of childcare demand, we estimate substantial welfare gains from the policy, yielding a Marginal Value of Public Funds (MVPF) above 3.5. Using the sufficient-statistic approach underestimates welfare gains by half. Counterfactual simulations and a difference-in-differences analysis suggest that increasing availability, rather than solely improving affordability, is crucial for the effective design of universal programs.

Keywords: universal childcare, daycare coverage, social welfare, sufficient statistics, non-marginal policy

JEL Codes: D61, H43, J13, J22

*We thank Matteo Bobba, Philippe Bontems, Marie Connolly, Nicole Fortin, Jean-François Fournel, Christian Hellwig, Jean-William Laliberté, Fabian Lange, Pierre Lefebvre, Thomas Lemieux, Philip Merrigan, Tímea Laura Molnár, Lea Nassal, François Poinas, as well as conference and seminar participants for valuable suggestions and comments. We further thank Loïc Courtemanche, Pascal Doray-Demers, Wendy Kei, and Chunling Fu for their support in the UBC and UQÀM Canadian Research Data Centres. Part of this research was conducted at the Québec Interuniversity Centre for Social Statistics (QICSS) and at the UBC RDC, part of the Canadian Research Data Centre Network (CRDCN). This service is provided through the support of QICSS' Member Universities, the province of Quebec, the Canadian Foundation for Innovation, the Canadian Institutes of Health Research, the Social Science and Humanity Research Council, the Fonds de Recherche du Québec, and Statistics Canada. All views expressed in this work and remaining errors are our own.

†*Corresponding author.* Department of Economics, University of Warwick (email: sebastien.montpetit@warwick.ac.uk).

‡Department of Economics, ESCP Business School (email: lcarrer@escp.eu).

§Department of Economics, University of British Columbia (email: pierreloup.beauregard@gmail.com).

1 Introduction

Over the past century, the rise in female labor supply—often referred to as a "quiet revolution" (Goldin, 2006)—has intensified the challenges for parents, particularly mothers, in balancing work and family duties. In response, many governments have proposed universal childcare programs to support maternal labor-force participation. Yet, while early-childhood programs targeted at disadvantaged families are found to yield substantial social returns (e.g. Heckman et al., 2010, 2013; List et al., 2021; Bailey et al., 2021; García et al., 2020, 2023), evidence on universal reforms is mixed (see Baker, 2011; Hotz and Wiswall, 2019; Duncan et al., 2023, for reviews). In addition, universal programs require substantial public expenditures. For example, the current Canadian national expansion has committed $30 billion in federal investments (Seward et al., 2023). Due to the high costs and uncertain benefits, it remains unknown whether universal childcare policies are socially desirable.

In this paper, we use the Marginal Value of Public Funds (MVPF) framework to assess the welfare impact of universal childcare provision in Québec in 1997, often deemed the most ambitious childcare reform in North America. The provincial government of Québec introduced low-fee, regulated childcare and simultaneously expanded daycare capacity to address supply shortages. As a result, the daycare coverage rate—measured as the number of available spaces per preschool-age child—tripled from 0.15 in 1997 to 0.45 in 2003.

Estimating the overall welfare effects of universal childcare programs poses two key challenges. First, standard sufficient statistics for welfare analysis, which assume marginal changes in fiscal policy, may be biased when applied to such reforms. Under the marginal-policy assumption, welfare benefits can be fully captured by reduced-form estimates of causal effects thanks to the envelope theorem, which guarantees the absence of first-order utility gains (Hendren, 2016). However, for non-marginal reforms, this method is valid only under strong linearity assumptions on labor-supply elasticities (Kleven, 2021). Second, beyond earnings gains, an in-kind policy like childcare provision affects households' welfare through multiple non-pecuniary channels. While parents likely enjoy spending time with their children, providing development-enhancing care can be exhausting (Chaparro et al., 2020). Increased availability reduces non-monetary costs of childcare use, such as commuting time and the effort required to find care when supply is limited.

Our empirical strategy explicitly addresses these two challenges. We calculate the policy's MVPF, defined as the ratio of beneficiaries' willingness to pay (WTP) for the reform to its net cost to the government (Finkelstein and Hendren, 2020). However, rather than assuming that the policy is marginal and thus has no first-order impacts on utility, we estimate beneficiaries' WTP for a non-marginal in-kind transfer using a model of childcare demand. Simulating the reform within our estimated model allows us to infer beneficiaries' WTP while accounting for the fact that they re-optimize their behavior in response to the policy. We then benchmark this structural estimate to the standard sufficient-statistic estimator.

In the first part of the paper, we estimate the reduced-form economic impacts on beneficiaries. First, we employ an intent-to-treat (ITT) difference-in-differences approach, comparing parents of young eligible children in Québec to parents of preschoolers in the rest of Canada. We confirm previous evidence of large positive impacts on maternal labor supply and childcare use both on the extensive and intensive margins, with no impacts on fathers (Baker et al., 2008). Then, using novel data on regional daycare coverage rates within Québec we manually digitized, we provide several pieces of evidence suggesting that local daycare expansions reflect a supply shock and are plausibly exogenous to parents' childcare demand. Exploiting this variation, we show that increases in maternal labor supply and childcare use are positively associated with local daycare expansion levels. We further show that these stronger responses in high-expansion areas are observed even among low-educated mothers, who experienced only small changes in prices after losing access to a means-tested childcare credit. Focusing on this subgroup allows us to hold the price of daycare constant, thereby isolating supply effects. Our findings thus suggest that the increase in availability, not just the decrease in prices, is an important channel of impact on childcare use and maternal labor supply.

Second, we examine the policy's long-run impact on eligible children later in life. We find that the negative effects on behavioral outcomes in childhood (Baker et al., 2008, 2019) do not translate into lower educational attainment or reduced earnings in their early career. As there is no fiscal impact on children and fathers, we focus on mothers when calculating welfare impacts.[1]

Finally, we quantify mothers' earnings gains from increased labor supply to obtain a benchmark value of the policy's benefits under the standard sufficient-statistic approach. Specifically, we use a non-linear difference-in-differences model to capture heterogeneity in earnings gains across the income distribution. In our welfare analysis, we consider these heterogeneous impacts to calculate the implied fiscal externality (i.e., the return to the government from increased maternal earnings).

In the second part of the paper, we estimate a structural model of childcare demand to infer mothers' WTP for the policy change, accounting for potential first-order impacts on utility and non-pecuniary benefits. Our model integrates non-monetary costs of childcare use into a setup where a mother decides how to allocate her time between work and childcare while considering her child's human capital accumulation. Building on previous models of childcare demand and maternal labor supply, our framework captures key trade-offs families face, such as balancing employment and care, and deciding how much parenting effort to exert at home.[2]

Our estimation algorithm leverages causal estimates from the reduced-form analysis to directly identify key model parameters. Specifically, we show how the Québec natural experiment can serve as an instrument for maternal care hours to identify a (potentially) non-linear cost of parenting effort. To assess the validity of our model, we then verify that it replicates the ITT impacts on maternal labor supply and childcare use. We simulate the policy's main features in our model, estimated using pre-reform data, and compare the simulated changes in mothers' choices to the reduced-form estimates. This comparison serves as an out-of-sample validation (see Todd and Wolpin, 2006, 2023). The model closely matches the key behavioral

[1]Behavioral problems could impact the government budget through other channels than income taxes. In robustness exercises, we calculate the potential costs associated with increased youth criminal activity documented in Baker et al. (2019), using recent estimates of the costs of crime that consider victimization costs and productivity losses. Given the nature of typical juvenile crimes, this additional societal cost turns out to be relatively small compared to mothers' gains in this context.

[2]See Ribar (1995), Blau and Hagy (1998), Del Boca et al. (2014), Griffen (2019), and Chaparro et al. (2020).

responses to the policy, supporting our structural assumptions on behavior.

In the last part of the paper, we build on our reduced-form and model estimates to assess the policy's welfare impact using both the sufficient-statistics and structural estimators. When considering only mothers' earnings gains, we find a benchmark MVPF of 1.40, meaning that an additional dollar of net government spending generates $1.40 in mothers' earnings. This result indicates that the policy is welfare-improving, though its MVPF is relatively modest compared to the MVPFs above 5 found for targeted preschool interventions (Hendren and Sprung-Keyser, 2020).

For the structural estimator, we use our model to estimate the WTP for the policy from the simulated reform. Specifically, we compute the equivalent variation—that is, the change in income under the status quo that would leave the mother indifferent between the policy and the status quo. This approach allows us to account for potential first-order impacts on utility and non-pecuniary benefits from a non-marginal policy. We find a WTP more than twice as large as when considering earnings gains alone. Our structural estimator yields an MVPF estimate of 3.56, indicating substantial social returns from universal preschool policies.

Finally, we perform counterfactual analyses to identify which features of the reform drive most of the welfare gains and provide insights on the optimal policy scheme. Specifically, we compare the WTP for simulated counterfactual reforms where *(i)* childcare prices remain at pre-reform levels, and *(ii)* there is no increase in daycare coverage. In the first counterfactual, we find that maintaining the higher pre-reform prices while increasing childcare supply reduces the WTP by only 16%. In contrast, in the second scenario, reducing prices without increasing coverage causes a much larger drop in WTP. This suggests that most of the welfare gains are due to the rise in daycare availability rather than the price reduction.

We then evaluate whether the government could have achieved even larger welfare gains under different policy schemes. Specifically, we compare the Québec childcare reform to alternative levels of price reductions and daycare expansions. Consistent with our previous results indicating that parents are willing to pay substantially more for increased coverage than for reduced prices, we find significantly higher MVPFs for reforms that invest more public funding into increasing availability rather than improving affordability. Thus, our results suggest that higher welfare gains could be achieved by allocating more resources towards creating additional daycare spots rather than lowering childcare fees.

This paper contributes to three strands of literature. First, we add to the large literature investigating impacts of universal childcare provision. Little is known about the overall societal implications of such policy changes. Most studies have focused on measuring the causal impact of preschool enrollment on child development or maternal labor supply, yielding mixed evidence. Effects on maternal labor supply largely depend on the counterfactual childcare market (Kline and Walters, 2016; Karademir et al., 2024).[3] Regarding child development, estimated impacts depend primarily on program care quality and socio-economic status (e.g. Havnes and Mogstad, 2015; Kottelenberg and Lehrer, 2017; Herbst, 2017; Felfe and Lalive, 2018; Cornelissen et al.,

[3] In contexts with a high prevalence of informal care (by grandparents, siblings, etc.), studies find little impact on parents' labor supply (Fitzpatrick, 2010; Havnes and Mogstad, 2011a,b; Kleven et al., 2024). Conversely, policies that crowd out parental care hours typically have positive impacts on maternal employment (e.g. Gelbach, 2002; Baker et al., 2008; Herbst, 2017; Hojman and Lopez Boo, 2022).

2018; Fort et al., 2020).[4]

Our work is the first to show that a universal childcare reform in a developed country can provide substantial returns on initial investment. We estimate mothers' WTP for the policy, including non-pecuniary gains, a dimension not addressed in previous cost-benefit analyses of universal childcare programs (e.g., Fortin et al. (2013); Haeck et al. (2018) for the same policy or Andresen and Havnes (2019) for Norway). Moreover, we demonstrate that negative short-run impacts on non-cognitive outcomes (Baker et al., 2008, 2019; Haeck et al., 2015) do not necessarily translate into depressed economic outcomes later in life. Using newly collected data, we also highlight the role of local daycare supply in shaping the impacts of universal programs, consistent with Yamaguchi et al. (2018) for Japan and Cornelissen et al. (2018) for Germany.

Second, our work contributes to the literature on empirical welfare analysis. Seminal papers in this literature show that, under standard assumptions, monetary gains are a sufficient statistic for beneficiaries' WTP in the case of marginal policy changes (Chetty, 2009; Hendren, 2016).[5] More recently, a body of work has theoretically identified conditions under which transparent sufficient statistics for non-marginal policy changes can be derived (Finkelstein and Hendren, 2020; Kleven, 2021). However, as Kleven (2021) argues, these conditions are often beyond empirical reach due to the complexity of expressing the welfare effect as a fiscal externality. Kang and Vasserman (2024) and Bergstrom et al. (2025) propose welfare bounds for non-marginal reforms, but these apply to rather specific policies, namely fiscal policy and notches, respectively.

We contribute by showing how the sufficient-statistic framework for welfare analysis can be fruitfully combined with structural approaches to make welfare statements about non-marginal policies. Our analysis indicates that for policies with substantial costs and significant non-pecuniary benefits, such as the Québec childcare reform, sufficient-statistic estimators can substantially underestimate welfare gains. This result has implications beyond childcare policy. Applying the sufficient-statistic framework to non-marginal reforms—a common practice as documented in our survey of MVPF estimates in the Policy Impacts Library—might significantly compromise welfare conclusions.[6] The sufficient-statistic approach is known to deliver only an approximation of welfare effects for non-marginal policy changes (Kleven, 2021). Our results provide empirical insights into the magnitude of this potential approximation error for the case of universal childcare. In our setting, this strategy underestimates the program's benefits by more than half.

Third, we contribute to a growing literature that combines reduced-form ex post estimation of policy impacts and structural modeling (see Todd and Wolpin, 2023; Buera et al., 2023, for reviews). While some recent studies, such as Griffen (2019) and Chaparro et al. (2020),

[4]A few studies, such as Guner et al. (2020), Daruich (2022), and Borowsky et al. (2022), estimate general-equilibrium models of the family to study impacts of childcare programs. Unlike those papers, we focus on an implemented reform rather than on a hypothetical policy scheme.

[5]Hendren (2016) highlights that this result applies to small changes in *fiscal* policy but that pecuniary benefits are no longer sufficient if the policy in question changes the state of public good provision.

[6]In Appendix C we provide a survey of MVPF estimates from the Policy Impacts Library of Hendren et al. (2025). Our (non-representative) sample of estimates suggest that computing the MVPF of non-marginal policy changes as if they were marginal is common practice. We stress, however, that this exercise requires several judgment calls and should only be seen as suggestive that policies considered in this literature are often non-marginal.

specify structural models of the family to interpret experimental impacts of targeted preschool programs (Head Start and IHDP, respectively), we do so in the context of universal childcare provision. Our work also relates to Chan and Liu (2018) who study a different policy scheme, which provided cash transfers to stay-at-home mothers in Norway. We contribute to this literature by showing that a tractable behavioral model, which incorporates non-pecuniary considerations of childcare decisions, can replicate the reduced-form impacts of the Québec program. Furthermore, we leverage the natural experiment to build a transparent identification argument for some key model parameters.

The rest of the paper is structured as follows. Sections 2 and 3 describe the institutional background and the data, respectively. In Section 4, we present our reduced-form analysis. Section 5 presents our structural model of childcare demand and its estimation. Section 6 presents our estimates of the policy's MVPF as well as counterfactual simulations. Last, Section 7 concludes.

2 Institutional background: The Québec childcare reform

On September 1, 1997, a large-scale reform of preschool daycare was initiated by the provincial government of Québec, the second most populous province in Canada. At the time, the province was lagging behind the other Canadian provinces in terms of female labor force participation. The major reform was thus designed to address this issue as well as to fight poverty and promote equality of opportunity for children (Japel et al., 2005). The centrepiece of the policy was the introduction of reduced-fare spaces in regulated childcare at an out-of-pocket price of $5 per day per child.

Implementation. The allocation of low-fee, regulated spaces was made through a decentralized process. The government incentivized the opening of childcare slots in a new network of non-for-profit, regulated facilities named *Centres de la petite enfance* (CPEs). Various forms of childcare could be part of the network, including family-based care (in the provider's home) and daycare centers. Interested families were added to a waitlist until a space in the desired facility became available. Providers had to follow their waitlist so they could not choose which family would get access first.

The reform was phased in by age of the child over a period of 4 years. In September 1997, only children aged 4 were eligible. In the following years, the age requirement was gradually lowered: one year later, 3-year-olds became eligible, followed by 2-year-olds in September 1999. In September 2000, subsidized spaces became available to all children aged less than 59 months. Access to the program was universal so that there were no entry requirements such as household income or participation to the labor force. Gross prices remained constant for parents until 2014—with the exception of an increase to $7 per day per child from 2004—with the provincial government subsidizing the remaining fees.

Net fee reduction. Prior to the policy, the average daily gross price was approximately \$21.[7] However, in the previous regime, families were eligible for a refundable childcare credit, which was decreasing with household income (see Appendix Figure A1).[8] After the reform, the daily fee was reduced to \$5 for all households. As a consequence, the reform was most beneficial for middle- and high-income households. Lefebvre et al. (2009) document that before the policy, the mean daily fee net of federal and provincial credits was approximately \$11 for middle-income and \$16 for high-income families. Therefore, considering a family signing a contract for the maximum number of days, this translates into annual savings of about \$1,500 to \$2,900, namely approximately 4% of household income. Conversely, the net price remained essentially the same for low-income families.

Expansion of the daycare market. The fee reduction was accompanied by a large expansion of the daycare market. There were important shortages before the reform, with only about one space per 6 preschoolers in March 1997. As such, despite a rapid expansion, there was still excess demand in the first years of implementation. There were long waiting lists at each regulated childcare facility. Figure 1, which shows the evolution of daycare coverage rates across regions within the Quebec province, illustrates this low supply. In 2000, only 29% of preschool-age children had access to low-fee services. This figure also reveals substantial heterogeneity in coverage rates within the province, which we exploit in our empirical analysis. For instance, in the pre-reform period, the region with the lowest coverage had only half the rate of the region with the highest coverage.

The slow growth in the number of spaces at the beginning is in large part due to the government's decision to freeze the number of spaces in for-profit, unregulated daycare. To remain as for-profit entities, daycare providers could only sign an agreement with the government and open additional spaces at a reduced fee. In practice, this moratorium induced nearly all private providers to convert to CPEs by 2002 (Haeck et al., 2016). In our empirical analysis, we focus on that period as we are interested in public childcare provision.

The moratorium on the creation of for-profit, unregulated daycares was lifted in 2003. New spaces kept being created at a fast pace over the following years, raising the share of children with access to subsidized spots above one half by 2005.[9]

Shortages before the reform. Despite evidence of substantial demand for childcare services before the policy,[10] the daycare market was characterized by very low supply. In the Québec context, these shortages are unlikely to be demand-driven by say, parental credit constraints (Caucutt and Lochner, 2020) because many high-educated (high-income) families were not using daycare. Several other reasons could explain why the private market failed to meet demand even if it was unregulated. It is widely known that government intervention is typically needed in

[7] See Office des services de garde à l'enfance (1997b).

[8] The reform also included the abolition, for households with a subsidized space, of some other universal family allowances. Baker et al. (2005) describe these other fiscal changes in detail and show that they have small impacts on the family's budget.

[9] Until 2009, growth in the regulated network was still superior to that in the for-profit network. Data assembled by Haeck et al. (2016), which we complemented with recent years using ministerial reports from 2017 to 2019, however, show that, from 2010, the for-profit network rapidly expanded as the regulated network stabilized.

[10] See Office des services de garde à l'enfance (1997a).

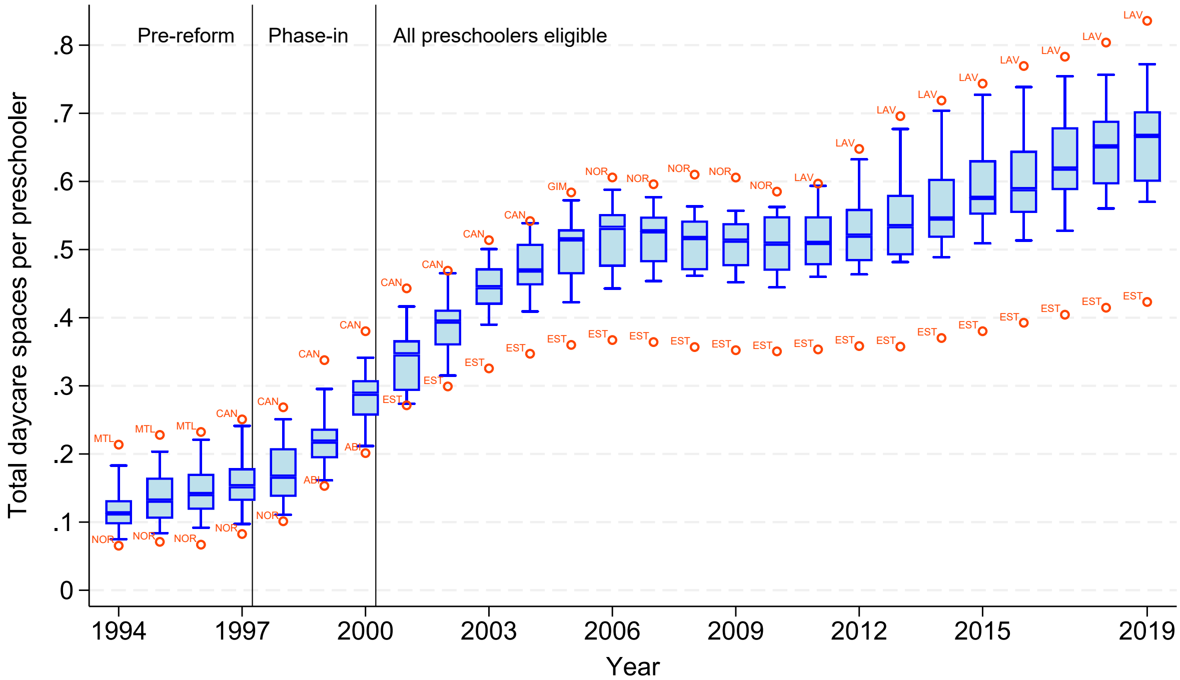


Figure 1: Evolution of the daycare coverage rate by administrative region, Québec

Note: This Figure displays the median (middle blue line), the 25th and 75th percentiles (light blue boxes), the 10th and 90th percentiles (blue brackets), and outliers (orange circles) of the daycare coverage rate distribution at the administrative-region level in a given year. It is constructed using data from a series of ministerial reports for daycare spaces (*Ministère de la Famille*, which held various other names), from the Québec Statistical Institute for the children population, and from authors' calculations. The dependent variable is the ratio between the total number of spaces (the sum of subsidized and non-subsidized spaces) for preschool children (as of March 31st) over the number of children aged 0-4 years (as of July 1st) in a given region. The reduced-fee program ($5/day/child) began in September 1997 only for children aged 4. All preschool children (0-4 years old) were eligible as of September 2000. In January 2004, the daily fee was raised to $7. ABI = Abitibi-Témiscamingue; CAN = Capitale-Nationale; EST = Estrie; GIM = Gaspésie-Îles-de-la-Madeleine; LAV = Laval; MTL = Montréal; NOR = Nord-du-Québec.

these markets. Even in the absence of government regulation, there are several natural barriers to entry. It is difficult for providers to signal quality, which might explain why quality in private settings is typically low (Cleveland et al., 2008). Moreover, operating in the childcare industry requires substantial capital (both human and physical), which may be difficult to obtain. In particular, retention of a highly qualified workforce is a challenge (Penn and Lloyd, 2012). The Québec government invested in these aspects by offering infrastructure subsidies, improving working conditions of childcare staff, and creating the CPE network as a signal of quality. Such investments might suggest that market imperfections were known by the government.

Lefebvre and Merrigan (2008) show that informal care by relatives was a small share of care arrangements used in Canada at the time. Most children in households using some form of care (around 41% in the pre-reform period) are in formal arrangements such as daycare centers or family-based care. Therefore, the relevant counterfactual care arrangement in our setting is care at home by parents (mothers).

Quality of care. Given the rapid expansion of the market, maintaining quality standards was a challenge. Nevertheless, an audit study conducted by Japel et al. (2005) found that, despite quality being low in CPEs, they still outperformed all other childcare settings on the vast majority of the criteria they considered for quality on average. Some efforts were also made by the government to gradually increase investments in quality in the subsidized network, such as raising educational requirements and wages of staff in 2000 over a four-year period (Molnár, 2023). Staff-to-child ratios remained largely unchanged despite the increase in maximum capacity (Baker et al., 2005). Appendix A.1 provides additional details on childcare quality following the reform.

3 Data sources

Our empirical analysis exploits several sources of Canadian micro-data. The main source is the National Longitudinal Survey of Children and Youth (NLSCY) of Statistics Canada (2010), which contains rich information on representative samples of Canadian children and their parents over the period of the reform. We notably observe measures of perceived care quality, labor-market participation of parents, and daycare expenditures, all of which are crucial to estimate parents' WTP for the policy. These biennial cross-sectional surveys covering the period 1994-1995 to 2008-2009 also contain a longitudinal component, allowing to follow a subset of children over several cycles and model human capital accumulation in early childhood.

Sample selection. As in Baker et al. (2008), we focus on two-parent families with a preschool age child because a simultaneous (but unrelated) change in Québec fiscal policy affected single parents. For the main analysis, we use survey years 1994-1995 and 1996-1997 as pre-reform data and the 2000-2001 and 2002-2003 surveys as post-reform periods. This choice is motivated by the institutional context. First, we omit the third survey wave (1998-1999) so as to focus on parents of children eligible over all of the preschool period. Doing so, we also mitigate concerns over treatment effect weighting in staggered designs and anticipatory behavior (De Chaisemartin and d'Haultfoeuille, 2020).[11] Second, we end our analysis in 2003 as some features of the reform evolved a few years after its implementation. In particular, daycare providers could open spaces in the unsubsidized, for-profit network from 2003 onward and the daily fee was raised to $7 in 2004. Since our goal is to analyze subsidized childcare provision, we focus on the reform as it was originally conceived.[12]

Table A1 reports summary statistics comparing Québec and the rest of Canada, our control group. In our analysis sample, we observe 34,042 children aged 0-4 and their parents.

Long-run analysis. To estimate the long-run effects of the childcare reform on eligible children's education, we use the Canadian Censuses of population of 2016 and 2021. These recent datasets allow us to compare individuals who are old enough to have completed their education.

[11] Ding et al. (2021) find suggestive evidence of strategic placement of children over the implementation period, especially by families with high maternal education. High-educated mothers were significantly more likely to pay for unsubsidized spaces to guarantee a subsidized spot once such new spaces would be open. This strategic response generated a disproportionate increase in childcare use in the province among younger children not yet eligible. Karademir et al. (2024) document similar anticipatory behaviors.

[12] Relatedly, a major change in Québec's parental leave policy occurred in 2006 and had a substantial impact on mothers' employment and earnings (see Patnaik, 2019; Karademir et al., 2024). For this reason, we do not use all cycles of the NLSCY, especially for the evaluation of the earnings impacts, as this other policy change might introduce bias in estimates of long-term effects.

For earnings in the early career, we use tax files data from the Longitudinal Administrative Databank (LAD) over the period 2010-2022. We relegate further details of these more standard datasets to Appendix A.2 and rather dedicate more space to describe our novel data source on daycare supply within Québec.

Daycare supply data. While previous studies of the Québec childcare reform estimate ITT effects of the policy at the provincial level, we investigate treatment effect heterogeneity at a more granular level *within* Québec. We assemble a novel dataset of the daycare coverage rate at the administrative-region level from annual management reports of the Ministry of the Family. These reports include information on the number of daycare spaces by administrative region by type of facility (centre-based, CPE, for-profit, etc.) from 1994 to 2019. Unfortunately, prior to 2004, we cannot distinguish between regulated and for-profit spaces. This is not a major issue for our empirical analysis for two reasons. First, the share of for-profit spaces in Québec was less than 7% of total spots until 2010, thus making the for-profit market rather marginal. Second, we restrict our empirical analysis to until 2002-2003, the period over which the government froze the number of for-profit spaces.

There are 17 administrative regions in Québec, which makes them a relatively granular level given that the total provincial population was approximately 7 million inhabitants in the late 1990s. In the Québec context, using the coverage rate at the administrative-region level is arguably preferable to more granular levels since many families send their children to daycare in other municipalities.[13]

4 Reduced-form impacts of the reform

We begin our empirical analysis by analyzing the reduced-form effects of the program on its main beneficiaries. First, we assess short-term impacts on parental time allocation (Section 4.1). Then, we turn to the long-run outcomes for children as they age, examining whether early exposure translated into long-run differences in educational and labor market outcomes (Section 4.2). Finally, we quantify the program's impact on earnings, which we later use to provide a benchmark estimate of the monetary value of the reform (Section 4.3).

4.1 Short-term impacts on parental time allocation

Average effects. We begin by summarizing the average effects of the reform on parents' time-allocation choices. Table 1 replicates estimates from prior studies, using the NLSCY sample defined in Baker et al. (2008) (henceforth BGM), based on the 1994–2003 survey waves. We focus on short-term impacts on maternal labor supply, childcare use, and parenting effort

[13]From other ministerial reports, we confirm that using daycare outside the place of residence is rare at the administrative-region level but is actually common at the municipal level. The share of children in daycare coming from other administrative regions is low (7.5% for Montréal and Laval and only 1.8% on average in other regions in 2001), but it is large at the municipal level. Indeed, in 2001, 25.1% and 15.8% of children in daycare in the Montréal and Laval regions were coming from another CLSC territory (a geographical unit grouping a few municipalities), respectively. A similar figure is observed in other regions: 23.6% of children in CPEs and 19% of those in other daycare came from other municipalities in 2001 (Ministère de la Famille et de l'Enfance, 2001).

(measured by the frequency of reading to the child).[14] Our empirical strategy uses an ITT difference-in-differences estimator, comparing two-parent families with preschool-aged children in Québec to similar families in the rest of Canada.[15]

Table 1: Average intent-to-treat (ITT) estimates

Dep. Var.:	Maternal employment		Childcare use		Reading to the child		
	Mother works	Mother's work hours	Child in care	Childcare hours	Rarely/ never reads	Reads weekly	Reads daily
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
β : Eligible$_{pt}$	0.078***	2.129***	0.138***	5.736***	-0.054***	0.044***	0.015
	(0.009)	(0.406)	(0.009)	(0.253)	(0.005)	(0.008)	(0.010)
	[0.029]	[0.132]	[0.003]	[0.001]	[0.028]	[0.140]	[0.616]
Mean dep. var.	0.532	17.54	0.418	13.07	0.226	0.748	0.379
R^2	0.105	0.099	0.116	0.110	0.170	0.053	0.165
N	33,758	33,637	33,709	30,915	33,171	33,171	33,171

Notes: This table summarizes intent-to-treat (ITT) estimates from difference-in-differences regressions assessing the impact of the Québec childcare reform. Each columns refers to a different outcome. We report estimates of the β coefficient from the following baseline specification estimated for mother i in province p in year t: $Y_{ipt} = \alpha + \beta \text{Elig}_{pt} + \gamma_p + \gamma_t + \delta X_{ipt} + \varepsilon_{ipt}$, where Elig_{pt} takes value 1 if the household resides in Québec after the reform. The outcome variable, Y_{ipt}, is maternal labor supply (extensive and intensive margins in Columns 1 and 2, respectively), childcare use (extensive and intensive margins in Columns 3 and 4, respectively), and parenting effort, measured as the frequency of reading to the child (rarely or never, at least weekly, and daily in Columns 5 to 7, respectively). X_{ipt} is a vector of controls including maternal age, child age, number of siblings, population of the area of residence, parental education, parents' immigration status, and provincial unemployment rates. γ_p and γ_t are province and survey year fixed effects. Standard errors clustered at the province level are displayed in parentheses. p-values computed using the Wild subcluster bootstrap of MacKinnon and Webb (2018) accounting for the small number of clusters are reported in brackets. The sample corresponds to the NLSCY sample following Baker et al. (2008), based on the 1994–2003 survey waves. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Consistent with previous results (see Baker et al., 2008; Molnár, 2023), we find that the reform substantially increased maternal employment, childcare use, and the propensity of reading to the child at least weekly by 7.8, 13.8, and 4.4 percentage points, respectively (corresponding to 14.7%, 33%, and 5.9% of pre-reform mean outcomes). The positive impacts also emerge at the intensive margin: mothers work an additional 2 hours per week on average, and childcare use increases by almost 6 hours. Table A2 further confirms that the increase in childcare use is driven by an increase in formal (institutional) care and that fathers' labor supply is unchanged.

Regional heterogeneity. The average impacts documented above capture a combination of the effects of the price decrease and the supply increase. To evaluate the role of supply, we now exploit our novel data on local daycare coverage. To this end, we first estimate an augmented version of our baseline specification in Table 1 that interacts eligibility with indicators for administrative regions within Québec, while maintaining the same baseline set of control variables and sample restrictions. The estimating equation is:

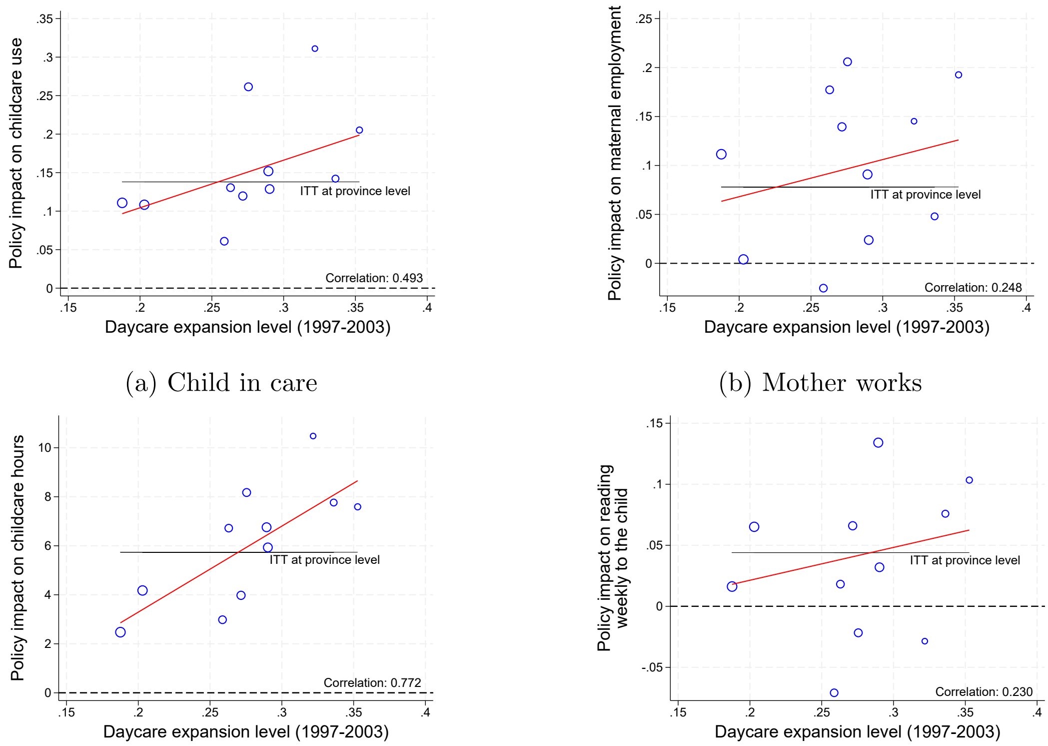
$$Y_{iprt} = \alpha + \beta_1 \text{Elig}_{pt} + \sum_r \beta_{2,r} \, \text{Elig}_{pt} \times \mathcal{I}_r + \gamma_p + \gamma_t + \gamma_r + \delta X_{ipt} + \varepsilon_{iprt} \quad (1)$$

[14]Our focus on short-term impacts is motivated by institutional factors discussed in Section 3.

[15]See the caption of Table 1 for further details.

where Y_{iprt} is either parental labor supply (extensive and intensive margins), childcare use (intensive and extensive margins), and the frequency of reading to the child, our measure of parenting effort. The eligibility indicator, Elig_{pt}, takes the value of 1 if the household resides in Québec after the reform. $\mathcal{I}_r$ takes the value of 1 if individual i lives in administrative region r within Québec.

In Figure 2, we plot the estimated policy effect for each region ($\beta_1 + \beta_{2,r}$) against that region's expansion in daycare coverage, measured as the difference in daycare coverage rates between 1997 and 2003.[16] We find a strong positive association between daycare expansion levels and policy impacts, especially for childcare take-up at both the extensive and intensive margins. For maternal employment and time reading to the child, the positive associations are less strong, but remain positive. This suggests that availability, and not only affordability, is a key driver of mothers' behavioral responses to the policy.



(a) Child in care

(b) Mother works

(c) Hours in care

(d) Reads weekly to child

Figure 2: Heterogeneous impacts of the Québec childcare reform and daycare expansion levels

Note: These figures plot program effects on (a) childcare take-up (any non-parental care), (b) maternal employment, (c) childcare hours, and (d) reading weekly to the child estimated using equation (1) against the childcare expansion rate at the administrative-region level. Circle sizes and the regression line are weighted by the number of observations. Included are regions with a sufficient sample size. The solid black line is the point estimate at the provincial level obtained from estimating our baseline specification, which we previously used to produce the ITT estimates in Table 1.

Heterogeneity by maternal education. To further isolate the role of daycare supply, we perform a heterogeneity analysis exploiting the differential financial benefits across family types. To finance the policy, the Québec government abolished a refundable childcare credit that was

[16]The full set of point estimates is reported in Appendix Table A3.

rapidly decreasing with household income. For low-income families, the difference in the net price of childcare introduced by the reform was thus very small. Focusing on low-income families allows us to hold the price of daycare constant, thereby isolating supply effects.

In Table 2, we use education as a proxy for income and estimate a simplified version of equation (1) separately for mothers who completed a post-secondary degree and those who did not. We group regions into two categories, namely high- and low-expansion regions. We define the low-expansion status as being in the bottom tercile of the distribution of daycare expansion over the period of analysis.[17] We then define treatment intensity by interacting the eligibility dummy with the low-expansion status.

Table 2: Heterogeneous impacts of the Québec childcare reform by daycare expansion and mother's education

Dep. var.:	Mother works				Childcare use			
Mother's education:	Low educ		High educ		Low educ		High educ	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β_1 : Eligible$_{pt}$	0.030	0.191***	0.095***	0.084*	0.075**	0.222***	0.158***	0.157***
	(0.019)	(0.029)	(0.079)	(0.045)	(0.014)	(0.078)	(0.027)	(0.016)
	[0.368]	[0.040]	[0.009]	[0.025]	[0.185]	[0.046]	[0.000]	[0.003]
β_2 : Eligible$_{pt}$		-0.099**		0.033		-0.082*		-0.01
$\times$LowExp$_r$		(0.047)		(0.044)		(0.042)		(0.037)
		[0.758]		[0.561]		[0.930]		[0.936]
Region (r) FE		✓		✓		✓		✓
r-level controls		✓		✓		✓		✓
p-value of $\beta_1 + \beta_2 = 0$		0.031		0.000		0.046		0.001
N	10,070	10,070	23,688	23,688	10,048	10,048	23,661	23,661
R^2	0.103	0.103	0.084	0.084	0.093	0.094	0.103	0.103

Note: This table presents estimates from difference-in-differences regressions assessing the impact of the Québec childcare reform by maternal education group. The high-education group is mothers who completed a post-secondary degree. In odd columns, we report estimates of the β_1 coefficient from the following baseline specification estimated for mother i in province p in year t: $Y_{ipt} = \alpha + \beta_1 \text{Eligible}_{pt} + \gamma_p + \gamma_t + \delta X_{ipt} + \varepsilon_{ipt}$, where Eligible$_{pt}$ takes value 1 if the household resides in Québec after the reform. In even columns, the estimating equation is: $Y_{ipt} = \alpha + \beta_1 \text{Eligible}_{pt} + 2EligibleptLowExpr + \gamma_p + \gamma_t + \gamma_r + \delta X_{ipt} + \varepsilon_{ipt}$. The impact in low-expansion regions is thus given by $\beta_1 + \beta_2$. X_{ipt} is a vector of controls including maternal age, child age, number of siblings, population of the area of residence, parental education and immigration status, and provincial unemployment rates. γ_p and γ_t are province and survey year fixed effects. Standard errors clustered at the province level are displayed in parentheses. p-values computed using the Wild subcluster bootstrap of MacKinnon and Webb (2018) accounting for the small number of clusters are reported in brackets. The sample corresponds to the NLSCY sample following Baker et al. (2008), based on the 1994–2003 survey waves. * p ¡ 0.10, ** p ¡ 0.05, *** p ¡ 0.01

Consistent with the literature, we find that the average employment impact is driven by high-educated mothers. The average impact on low-educated mothers is small and statistically insignificant. However, introducing heterogeneity by local daycare supply reveals that in higher-expansion regions, low-educated mothers do significantly increase their labor supply. This estimated impact is twice as large as that of high-educated mothers in the same regions. Moreover, we find no statistically significant difference in the impact of the policy by coverage status among high-educated mothers. These results are consistent with the financial incentives mentioned above: for high-educated mothers, our results suggest the main incentive to take-up

[17] This choice is motivated by the distribution of expansion levels across regions: as shown in Appendix Figure A2, many observations are concentrated around the median, so comparing regions below and above the median would contrast areas with very similar levels of exposure. Conversely, our approach ensures a clearer separation between low- and high-expansion areas.

employment was the price reduction; for low-educated mothers, access to a space was key. This also shows up in childcare take-up, where the stronger response in high-expansion regions is again driven by low-educated mothers.

Taken together, our findings suggest that the relief of capacity constraints in daycare at the local level was an important driver of the policy's impacts on mothers' time allocation. In other words, not only the price decrease but also the increase in capacity at the local level was an incentive for mothers to take up employment and use childcare.

4.1.1 Robustness checks

Alternative specifications. In Appendix A.4, we consider an alternative specification to examine treatment effect heterogeneity across regions with different levels of expansion, using the definition of low- versus high-expansion regions defined above. This specification yields the same qualitative results: the policy impacts on our main outcomes of interest (childcare utilization, maternal labor supply, and reading time) are driven by regions that experienced larger expansions in daycare supply. We also show that results are robust to using only Ontario—the province most similar to Québec—as a control group (Appendix Table A4).

In addition, to ensure that the results are not driven by differential trends in rural and urban areas within administrative regions, we estimate a version of equation (1) that includes interactions between survey wave dummies and an indicator for rural municipalities. The results, reported in Appendix Figure A3, remain very similar.

Absence of pre-treatment trends. As with any difference-in-differences strategy, a key threat to identification of the policy's treatment effects is the possibility of differential trends between the treatment and control groups prior to the reform. Many papers on the Québec childcare program have provided robust evidence that Québec and the rest of Canada (RofC) were following similar trends on a wide variety of outcomes prior to treatment (e.g. Baker et al., 2008, 2019; Haeck et al., 2015, 2018; Molnár, 2023). However, we might be concerned that regions within Québec were evolving differently prior to the policy. While there is no direct test of the parallel trends assumption, Figure 3 provides graphical evidence that Québec high-expansion regions, Québec low-expansion regions, and the rest of Canada were on similar trends for our outcome variables before the reform.

Appendix Figure A4, which reports estimated coefficients from event-study regressions, further confirms the graphical analysis. These regression results show that prior to the policy, the mean outcomes of interest were either converging or not statistically different before sharply diverging in post-policy waves.

Lastly, in Appendix Table A5, we show that pre-reform characteristics in low- versus high-expansion regions are very similar. Not only are families in the two groups of regions comparable in terms of demographic characteristics (such as parents' education and the number of children in the household), but also along our outcomes of interest. This evidence strengthens our confidence that low-expansion regions were not following differential trends.

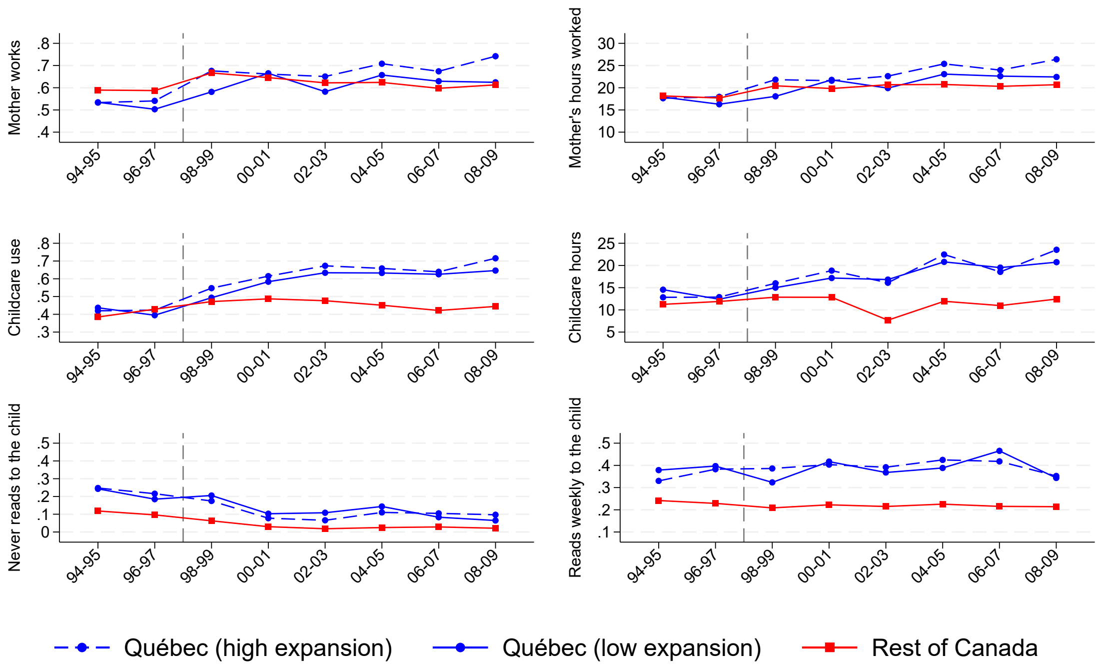


Figure 3: Mean childcare use, maternal labor supply, and reading time by daycare expansion status (in Québec) and in the rest of Canada

Note: These figures plot the means of selected outcome variables separately for three groups, namely two groups within Québec (in blue) and the rest of Canada (in red) in the NLSCY. The dotted (resp. solid) blue line represents families living in regions where the expansion of daycare coverage falls into (resp. is above) the bottom tercile in the province over our study period. The data source is the 8 waves of the NLSCY. The sample is restricted to two-parent families with a preschool-age child.

Exogeneity of local childcare expansions. To interpret the evidence of heterogeneous responses as causal, local childcare expansions must be uncorrelated with the error term in our regressions. In particular, this means that expansions should reflect a change in daycare supply rather than being driven by increased demand. We present evidence below to support the reliability of this assumption.

First, we explore the potential determinants of changes in regional coverage rates (see, e.g. Cornelissen et al., 2018; Yamaguchi et al., 2018). In Figure 4, we plot the correlation between the daycare expansion levels and pre-reform region-level characteristics. Reassuringly, the figure reveals that virtually all the considered characteristics are uncorrelated with expansion levels. In particular, in the last panel, we note that there is no systematic correlation between regional pre-reform childcare prices and expansion levels, helping us rule out the concern that lower pre-reform prices reflect lower demand and mechanically smaller impacts of the reform. There are only two exceptions: the initial coverage rate and the share of high-educated individuals. While the negative correlation with initial coverage is mostly mechanical, the correlation with education levels raises the concern that local expansions may reflect differences in educational attainment. To address this, we control for education shares in some specifications in the following section, along with region fixed effects, which absorb differences in time-invariant regional characteristics. As additional evidence to support our main identification assumption, in Appendix Table A6, we regress the expansion level on all considered characteristics. Reassuringly, we find that none of these variables can predict local childcare expansions. Indeed, we cannot reject the null hypothesis that the estimated coefficients are jointly zero.

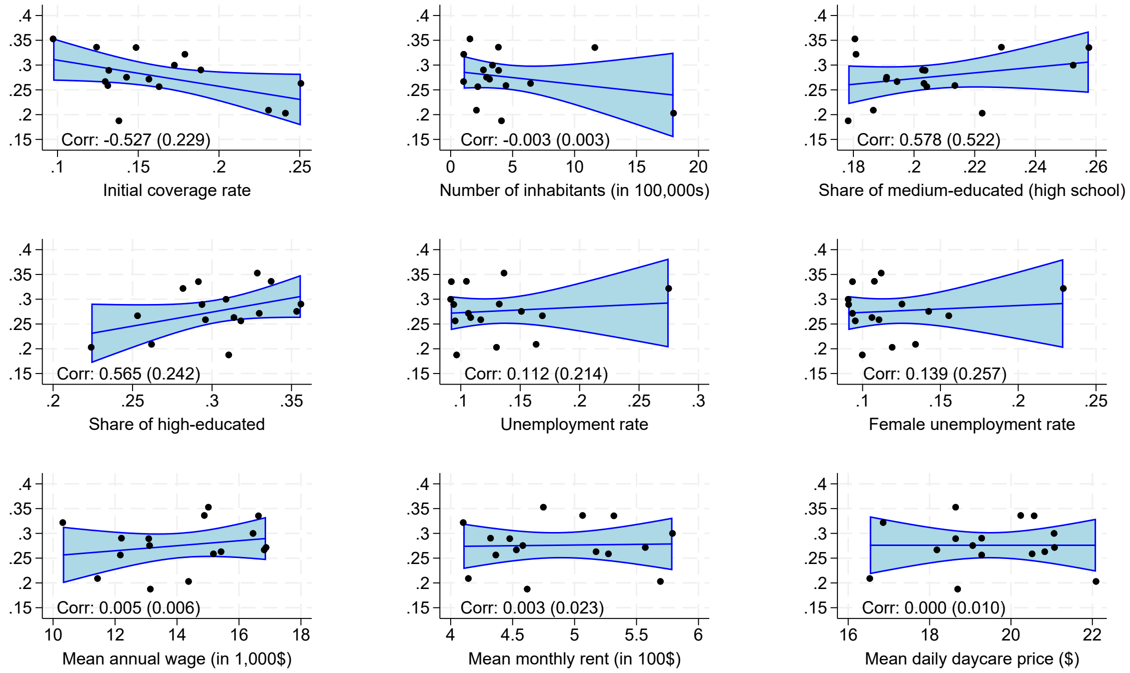


Figure 4: Correlations between regional daycare expansion and region-level characteristics

Note: These figures illustrate the relationship between the daycare expansion level (y-axis) and pre-reform region-level characteristics potentially associated with childcare demand (as indicated on the x-axis). The daycare expansion level is defined as the increase in the daycare coverage rate from 1997 to 2003. Region-level variables are calculated from 1997 data of the Institut de la Statistique du Québec (2023) and the 1996 Canadian Census of population. Shaded areas are 95% confidence intervals. Robust standard errors on univariate regression coefficients are reported in parentheses.

A second potential threat to identification could arise from households endogenously sorting into different regions according to childcare availability. To get a sense of whether residential sorting is an issue in our setting, we study the trends in interregional migration of families with a preschool-age child in Appendix Figure A5. This analysis reveals similar migration flows to regions in the top and bottom terciles of the daycare coverage distribution. Therefore, it is not the case that families systematically moved to areas that experienced greater increases in daycare supply.[18]

4.2 Long-term impacts on children

Having established that the policy has significant impacts on mothers' labor-market behavior, we now investigate long-run effects on eligible children's educational attainment and earnings as they age.[19]

[18]Another plausible demand-side channel could be that mothers take-up employment in public childcare services. However, data from the 1996 and 2001 Canadian Census suggests that the share of mothers' employment in our subpopulation of interest (mothers of preschoolers in two-parent families) is very small and in fact does not increase from 1996 to 2001. This share rather decreases from 3.69% in 1996 to 1.5% in 2001.

[19]The data, empirical strategy, and results are detailed in Appendix A.5.

For schooling outcomes, we use the Canadian Censuses of 2016 and 2021 and we implement a triple-difference estimator, which compares same-age individuals who vary in eligibility status based on the census year and their province of birth. In short, we find no evidence of negative effects on educational attainment of eligible children in the long-run. This pattern holds for each educational level, namely university, high school, and college/CEGEP completion.

We then use tax files data from the Longitudinal Administrative Databank to investigate long-run impacts on earnings. Using an event-study specification, we find insignificant positive effects on eligible cohorts' labor income and net earnings.

These findings suggest that we can rule out large negative effects on children's long-run economic outcomes, despite the rapid expansion of the childcare network. Previous research documented short-term negative effects of the reform on non-cognitive behavioral and health outcomes (Baker et al., 2008; Haeck et al., 2015), and possible persistence later in life (Baker et al., 2019).[20] These studies, however, find no systematic impact on cognitive skills. Moreover, parents compensated for formal care by increasing time and monetary investments in their children at home (Molnár, 2023), which may also help explain the absence of long-run economic penalties.

Given the absence of significant economic impacts for fathers and children, the remainder of the paper focuses on quantifying mothers' benefits.

4.3 Impacts on mothers' earnings

We now turn to analyzing the impact of the reform on mothers' earnings, as they constitute the main source of fiscal externalities and provide a benchmark value of the reform's benefits under the sufficient-statistic approach. We begin by estimating the average impact using our baseline specification, which we previously used to produce the ITT estimates in Table 1, with mothers' annual labor earnings from the NLSCY as the outcome.

Because earnings gains have different fiscal impacts along the income distribution of mothers, we move beyond average impacts and investigate how the policy shifts the income distribution. To estimate impacts on income in an absolute (unconditional) sense, we use the unconditional quantile regression framework of Firpo et al. (2009). This approach estimates quantile treatment effects by comparing the cumulative earnings distributions in pre- and post-reform periods in treatment and control groups using a recentered influence function (RIF) regression. Point estimates then indicate by how much a given quantile of mothers' labor income in Québec has shifted due to the policy.

Results of these two exercises are reported in Figure 5. We find that, on average, mothers of preschoolers in Québec earn an additional $3,750 (in constant 1997 dollars) per year in the post-reform period compared to other Canadian mothers of young children. Despite using a different dataset and focusing on a different age group than Lefebvre and Merrigan (2008), we reassuringly obtain a point estimate that is quantitatively comparable to their result.[21] The

[20]We note that Haeck et al. (2018) challenge some of these findings, showing that accounting for treatment dosage reduces the magnitude of long-term effects.

[21]Using the Survey of Labour Income Dynamics (SLID) covering the period 1993-2002, Lefebvre and Merrigan (2008) estimate a positive impact of 2,486$ (once adjusting to 1997 dollars) on mothers of children aged 1-5. Apart from the different sources and age ranges considered, the difference in the estimates might stem from the

quantile analysis reveals that there is a positive effect of about \$1,100 at the 4th and 9th deciles and larger impacts of \$4 to \$7 thousands in between.

To assess the plausibility of the parallel-trends assumption, we also estimate an event-study regression. Appendix Figure A6 confirms that earnings of Québec mothers are not statistically different in the pre-reform waves of the NLSCY, but start diverging only in the post-policy periods.

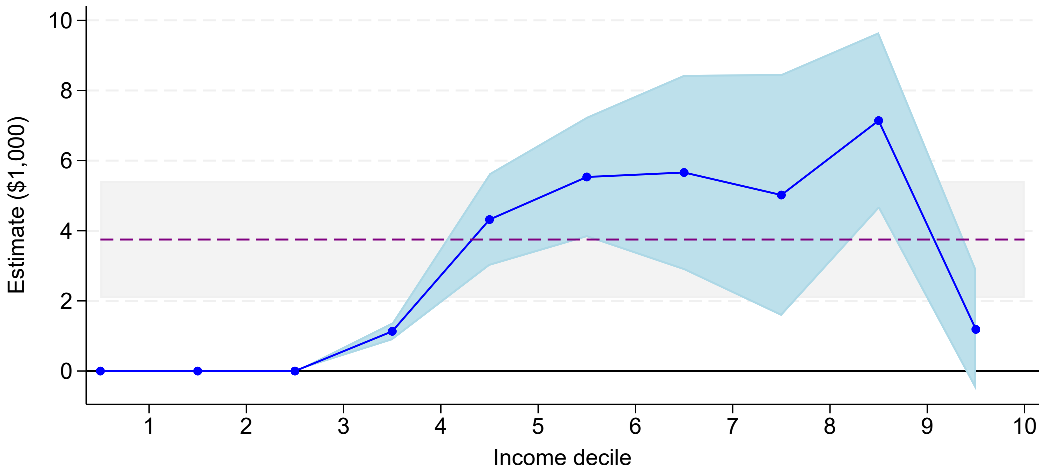


Figure 5: Distributional impact of the Québec childcare reform on mothers' labor earnings

Note: This figure displays the mean impact (dashed purple line) of the Québec childcare reform along with point estimates from RIF unconditional quantile regressions (blue line) of the policy impact on mothers' annual labor earnings (in constant 1997 thousand dollars). The sample is restricted to two-parent families with a preschool-age child and with non-missing covariates. Shaded areas are 95 percent confidence intervals around each estimate.

In our analysis of the policy's fiscal externality in Section 6, we take into account the impact of the documented heterogeneity of earnings gains on the government's budget. Before moving to estimating the MVPF of the Québec reform, we describe the economic model used to infer mothers' willingness-to-pay.

5 An empirical model of childcare demand

For our preferred MVPF estimator, we use a structural model of behavior to account for parents' behavioral responses and to quantify non-pecuniary gains. We consider a model of the family which builds on Chaparro et al. (2020) (henceforth CSW). Our main departure from CSW is that we do not assume that families have access to a full menu of childcare options. Motivated by the evidence of shortages in daycare presented in Section 2, we rather consider that families have access to only one childcare option.[22] To account for non-monetary costs of daycare, we

inclusion by the authors of 1999, where earnings gains are small and statistically insignificant, as a post-reform period. We refrain from including implementation years in our analysis due to issues with treatment effect weighting in staggered designs and anticipatory behavior (De Chaisemartin and d'Haultfoeuille, 2020).

[22]As documented in Section 2, childcare markets in Québec have been characterized by important shortages for several decades. Thus, assuming that childcare is available at any quality (and associated price) to every

incorporate local coverage into the household's decision problem leveraging our novel data on regional supply.

After describing the model in Sections 5.1 and 5.2, we briefly explain the numerical procedure to solve it in Section 5.3. We then discuss identification in greater detail in Section 5.4. Our strategy exploits the natural experiment generated by the Québec reform to identify some key model parameters. Finally, we present estimation results and model fit in Section 5.5.

5.1 Setup

The model is that of a time-allocation problem of a mother (a unitary household) with a young child (after the parental leave period) that has to meet her child's care needs. The mother weighs the consequences of her choices on the child's development, the family budget, and her own preferences, thus providing a framework that highlights the key trade-offs families face (Becker, 1965, 1991; Del Boca et al., 2014; Chaparro et al., 2020; Berlinski et al., 2024).

Time constraints. A mother of (at least) one preschool-age child makes a static decision on how to allocate her time T between market work L, child care at home T_m, and leisure ℓ. The mother's weekly time budget is given by:

$$T = L + \ell + T_m \tag{2}$$

Taking as given her child's baseline skills, denoted h_0, as well as household characteristics X, she jointly decides how to meet a child's care time constraint. While the child is awake (T_c hours), they must be cared for by the mother or in non-maternal care and thus we have:

$$T_c = T_m + T_d \tag{3}$$

where T_d is hours of non-maternal-care.[23]

Budget constraint. Each hour devoted to labor-market work is remunerated at a wage rate w. To finance non-maternal care expenses and household consumption C (there is no savings), she can use her own labor income as well as non-labor income I, which includes labor income from a spouse. The mother's budget constraint writes:

$$C + (1 - \tau_d(wL + I))\, p\, T_d = wL + I \tag{4}$$

where p is the hourly price of non-maternal care. The household receives refundable provincial childcare credits that account for a share τ_d of childcare expenses. The function τ_d is decreasing in household income, from 0.75 for the poorest households to 0.26 for households with income above \$48,000 (see Appendix Figure A1 for a graph of this function).

household as in CSW appears unreasonable. Moreover, in the Canadian context, daycare prices are not a strong predictor of quality (Cleveland and Krashinsky, 2009; Seward et al., 2023).

[23]Non-maternal care includes hours in daycare centres, family-based daycare, and care by relatives. We therefore assume that care by relatives is not free. This assumption is often made in the literature because it is otherwise difficult to rationalize not taking-up free care by a relative (e.g. Bernal, 2008; Bernal and Keane, 2010; Apps et al., 2016; Griffen, 2019; Guner et al., 2020). Rather than assuming an *ad hoc* process for how informal care might be available to some families and not others, we follow this stream of literature.

Child skill production technology. In addition to choosing a time allocation (L, T_m, ℓ, T_d), the mother chooses *how* to parent. She cares about her child's achievements, which are determined through a child-development production function H taking as inputs the child's initial skills h_0 as well as time and quality of care in each care mode. We also allow the child's skills at the end of the preschool period h_1 (measured at ages 4-6) to depend on a vector of household characteristics X such as the number and ages of other children in the household and the parents' education.

$$h_1 = H(T_m, T_d, q, e; h_0, X) \tag{5}$$

where q is the quality of non-maternal care, which varies across households, and $e > 0$ is the effort devoted by the mother into adopting parenting practices that stimulate the child's development.

Access to proximity childcare. The mother receives one offer of childcare (p, q), where p and q are the price and quality of non-maternal care, respectively. Searching for childcare entails a cost $\gamma_{d,1}$. With exogenous probability π, the search is successful and the mother finds a spot in her region of residence. If the search is not successful, the mother only has access to a spot far away, in which case she incurs a travel cost $\gamma_{d,2}$ if she uses non-maternal care.[24] Therefore, upon entering the childcare market, the mother incurs a non-monetary disutility, which we denote $\psi(T_d)$ and takes the form:

$$\psi(T_d) = \mathbb{1}[T_d > 0]\ (\pi\ \gamma_{d,1} + (1-\pi)(\gamma_{d,1} + \gamma_{d,2})) \tag{6}$$

This non-monetary cost is assumed to be fixed, meaning it does not depend on hours spent in non-maternal care.[25] We further assume that the probability to find a space in the region of residence is the same for every mother, so that π is equal to the regional daycare coverage rate.

Measurement. To proxy for e, we use reading time to the child.[26] To measure non-maternal care quality, we sum indicators of parents' perception about the frequency of caregiver-child interactions, the caregiver praising the child, and activities that stimulate learning, each taking values from 1 (never) to 4 (often). As for daycare expenses, those variables are only measured for individuals using childcare and are only available in post-reform waves of the NLSCY.[27] Therefore, for each household, we impute non-maternal care quality and the hourly daycare price using characteristics common to all waves (see Molnár, 2023). The variables used for

[24] Several recent studies show that families are sensitive to distance to caregivers. See Borowsky (2019); Bravo et al. (2022); Bodéré (2023); De Groote and Rho (2024).

[25] This modeling choice is consistent with our context given that childcare providers typically opened full-time spaces only (Haeck et al., 2018).

[26] Compared to Chaparro et al. (2020), who treat effort as unobserved and measure "quality of care", we instead use our data on parenting practices as a proxy of the former. CSW use items of the Home Observation for Measuring the Environment (HOME) index measuring parental Support for Learning and Literacy as their measure of maternal care quality. We do not have such a variable in the NLSCY data, so we proxy for parenting effort using the frequency of reading to the child.

[27] We note that measures of non-maternal care quality are observed in waves 3-4 of the NLSCY (1998-1999 and 2000-2001). Therefore, those variables are arguably not yet affected by the increased investments in quality by the Québec government initiated in 2000.

prediction are the age and number of siblings in the household, parents' age, education and immigration status, and the size of the area of residence. Details of these predictions are presented in Appendix A.6.

The decision problem. Mothers' utility depends on household consumption, time and effort parenting the child, leisure time, and the child's skill accumulation at school age (4-6 years old). Additionally, mothers' preferences include a non-monetary disutility of childcare use (when $T_d > 0$), intended to capture travel time to the childcare provider or search effort to find a spot. The mother's decision problem is to choose a time allocation (ℓ, L, T_m, T_d), a level of effort e, and a consumption level C to maximize her utility under the four constraints described above. Formally, the decision problem of a mother is:

$$\underset{\Gamma}{Max} \quad U(C, \ell, h_1, T_m, T_d, e) \quad s.t.\ (2), (3), (4), (5) \tag{7}$$

where $\Gamma \equiv (C, L, \ell, T_m, T_d, e)$ is the vector of choices.

5.2 Functional forms

Preferences. The mother's (household's) utility function is given by:

$$U = \gamma_c \ln(C) + \gamma_\ell \ln(\ell) + \gamma_h \ln(h_1) + \gamma_m \ln(T_m) - \gamma_{e,1}\, e\, T_m^{\gamma_{e,2}} - \psi(T_d) + \varepsilon \tag{8}$$

where ε is the unobserved component of utility and $\psi(T_d)$ is the non-monetary disutility of childcare use (defined above). A parameter of particular interest in this equation is $\gamma_{e,2}$, the non-linearity in the disutility of parenting effort, which in CSW's terms captures an "exhaustion effect" in maternal care. This feature of CSW's model is particularly relevant in our context since reducing exhausting parenting time represents an important source of non-pecuniary utility gain for parents.

Child skills production function. The child accumulates human capital in the preschool period based on the care received and initial skills h_0. Child skill at age of school entry denoted h_1 is given by:

$$\ln\,(h_1) = \delta_0 \ln\,(h_0) + \delta_e \frac{T_m}{T_c} \ln(e) + \delta_d \frac{T_d}{T_c} \ln(q) + X'\delta_m + \eta_h \tag{9}$$

where η_h is a productivity shock and the δ parameters capture the productivity of various inputs in generating skills. In this specification, quality of care in each care mode is weighted by the share of time the child spends in it so as to ensure that a given care mode impacts the child's development only when the child is actually there.

5.3 Model solution

Given the potential presence (and importance) of corner solutions in the model, it has to be solved numerically. Combining the budget and time constraints (2), (3), and (4) and plugging them into the objective, we are left with three choice variables. That is, the mother chooses a time allocation (L, T_d) and a level of parenting effort e. The solution algorithm works as follows. We build a grid over the feasible time allocations (applying the time constraints) and the effort level. Then, for each combination of time and effort on the grid, we compute the utility level using equation (8) and find the vector yielding the highest utility on the grid.

5.4 Identification and estimation

We adopt a transparent multi-step identification strategy following CSW. The advantage over a joint estimation algorithm is that it better isolates the sources of variation in the data we use for identification of key parameters of the model.

The first steps consist in identifying a set of parameters that do not require additional structure. These are parameters governing the child skills production function (9) and the exhaustion effect $\gamma_{e,2}$. Taking these productivity parameters as given, we then estimate the remaining preference parameters using a logit specification.

5.4.1 Child skill production technology

We first consider identification of the productivity parameters of the child skill technology (δ). We observe the time allocation of the child across different care modes as well as proxies for care quality as perceived by the person most knowledgeable about the child. Her parenting practices and household characteristics are also observed. In our baseline model, we estimate equation (9) by OLS using our measures of child development. We include a set of control variables that might influence child development such as the age and number of siblings in the household and parents' characteristics (age, education, and immigrant status).

In this step, we leverage the panel dimension of the NLSCY, which has rarely been used in the literature, to obtain a measure of h_1 using the children observed longitudinally. We use test scores in the Peabody Picture and Vocabulary Test (PPVT) conducted during the home interview as a measure of endline skills. Development at early ages h_0 is measured using the Motor and Social Development (MSD) score. This MSD score is constructed by Statistics Canada using a series of questions on dimensions of the motor, social, and non-cognitive development of young children. The two indices are standardized by age in months and by 2-month age groups, each with mean of 100 and standard deviation of 15.

The baseline OLS estimator might be biased since childcare choices are potentially endogenous. Several variables in the error term η_h such as the mother's innate parenting ability might be correlated with childcare decisions and child development. Also, quality of care in each mode is measured with error.

To address these identification issues, we consider an instrumental-variable approach leveraging the policy change to identify δ_e and δ_d. Let $\tilde{X}_i \equiv [\frac{T_m}{T_c}\ \ln(e),\ \frac{T_d}{T_c}\ \ln(q)]$ denote the row vector of endogenous variables in (9). As candidate instruments for $\tilde{X}_i$, we consider the eligibility dummy Elig_{pt} (taking the value of one if the family lives in Québec post-reform) and its interaction with low-expansion status (bottom tercile of expansion distribution). Sufficient conditions for identification are the usual instruments' relevance and exclusion-restriction conditions. The first identification assumption is that the instruments $Z_i \equiv [\text{Elig}_{pt},\ \text{Elig}_{pt} \times \text{LowExp}_r]$ are correlated with $\tilde{X}_i$. The heterogeneous impacts documented in

Section 4.1 lend support for this assumption: the reform induced a large increase in childcare use and maternal employment, especially so in regions with higher daycare coverage. Second, the exclusion restriction requires that Z_i is uncorrelated with the error term in equation (9). The identification assumption is thus that, conditional on initial skills and household characteristics, the policy should impact child development only through childcare choices.

5.4.2 Identifying the exhaustion effect $\gamma_{e,2}$

We next consider identification of the curvature of the (dis)utility of effort $\gamma_{e,2}$. In CSW, this exhaustion effect is identified directly from reduced-form estimates of the IHDP experiment thanks to random treatment assignment. Our natural experiment gives us similar variation to exploit for identification. Specifically, we leverage the optimality condition for the effort choice and our reduced-form estimates of the policy's impacts to identify $\gamma_{e,2}$. The optimal effort level, which must be at an interior solution (it is not defined at 0), is given by the first-order condition of the maximization problem with respect to e:

$$e^* = \frac{\gamma_{h_1}}{\gamma_{e,1}} \, \delta_e \, \frac{T_m^{1-\gamma_{e,2}}}{T_c} \tag{10}$$

Taking logs on both sides yields:

$$\ln(e) = (1 - \gamma_{e,2}) \ln(T_m) + \chi \tag{11}$$

where $\chi = \ln(\gamma_{h_1}) - \ln(\gamma_{e,1}) + \ln(\delta_e(h_0)) - \ln(T_c)$. Thus, optimal (log) parenting effort is determined by maternal care time (T_m) and a sum of productivity and preference parameters. As CSW note, a simple OLS estimator of $(1 - \gamma_{e,2})$ would be biased because T_m is likely correlated with preference and productivity parameters in χ. We therefore consider using the quasi-experimental variation to identify the exhaustion effect.

Assuming that potential average *changes* in χ conditional on individual characteristics X are the same in Québec and the rest of Canada, we can leverage our difference-in-differences estimates to identify $\gamma_{e,2}$. Given the evidence supporting the parallel-trends assumption (conditional on household characteristics X) for parenting effort and maternal-care time, our DiD estimates identify the Intent-to-Treat effects on those outcomes. Thus, conditioning equation (11) on X as well as province ($Q = 1$ for Québec) and a post-treatment dummy and then taking the double-difference yields:

$$\begin{aligned} &\text{ITT}[\ln(e)] = (1 - \gamma_{e,2}) \, \text{ITT}[\ln(T_m)] \\ &+ \mathbb{E}[\chi|Q = 1, post, X] \; - \; \mathbb{E}[\chi|Q = 1, pre, X] - (\mathbb{E}[\chi|Q = 0, post, X] \; - \; \mathbb{E}[\chi|Q = 0, pre, X]) \end{aligned} \tag{12}$$

where ITT$[A]$ is the intent-to-treat impact on variable A. Therefore, assuming that the evolution in χ (conditional on individual characteristics) is the same in Québec and the rest of Canada, the second line in equation (12) is null and the ratio of treatment effects on parenting practices and time identifies $\gamma_{e,2}$. This assumption is plausible in our context given the evidence in the literature of similar trends between the two groups over a wide range of outcomes.

To lend some additional support for this identification assumption, in Appendix Table A7 we report estimates of the child skill production technology in different Canadian provinces. Reassuringly, we cannot reject the hypothesis that the productivity of parenting δ_e in Ontario, British Columbia, the Prairies, and the Atlantic provinces is the same as that in Québec. Parenting productivity thus should not had evolved differently in Québec compared to our control group, the rest of the country.

5.4.3 Preferences

Taking as given the primitives estimated in the previous steps, we estimate preference parameters on pre-reform data using the Québec sample only. We discretize work hours (L), hours in nonmaternal care (T_d), and reading time (e) to obtain a discrete-choice representation. We then assume that the unobserved component of utility ε follows an i.i.d. Gumbel distribution, which yields a standard conditional logit model for preferences (McFadden, 1974). This distribution for the unobserved component of utility has the well-known advantage of yielding a closed-form expression for choice probabilities. We estimate the preference vector $\gamma \equiv (\gamma_C, \gamma_\ell, \gamma_{h_1}, \gamma_{T_m}, \gamma_{e,1}, \gamma_{d,1}, \gamma_{d,2})$ by maximum likelihood.

5.5 Estimation results

Before computing our MVPF estimates using the model in the next section, we begin by presenting the estimates of the model's main components. We then discuss the model fit and contrast the policy impacts estimated in the reduced form to the predictions of the model.

5.5.1 Model parameters

The first two panels of Table 3 reports model parameters estimated in the first two steps, namely the exhaustion effect ($\gamma_{e,2}$) and productivity parameters (δ). For both specifications, we compare OLS and instrumental-variable models leveraging variation from the policy change.

Child human capital. We first report the point estimates for the key inputs of the child human capital production function. These are the productivity of initial skills (δ_0), maternal care (δ_e), and non-maternal care (δ_d). Unsurprisingly, those three inputs are indeed productive of child human capital. Consistent with previous literature on child development, we find that early-age skills are highly predictive of future skills (see Cunha and Heckman, 2007). The OLS estimate (column 1) suggest that a 10% increase in the Motor-Social Development Score translates into a 2.03% higher PPVT score at ages 4-6.

Childcare time and quality are also positively associated with endline child skills. In the linear regression (column 1), we find that a 1% increase in parenting quality per hour in maternal care is associated with an increase in the child's PPVT score of 0.154%. Similarly, a 1% higher quality per hour in non-maternal care increases the child's endline skills by 0.210%. Those parameters, however, should be interpreted with caution because movements in one input involve manipulating several endogenous variables: an increase in maternal-care time (T_m) implies a reduction of non-maternal care (T_d). Additionally, reading time (e) might depend on T_m through the exhaustion effect.

These baseline OLS estimates might suffer from omitted-variable bias through, for example, some unobserved innate parenting ability, which is correlated with childcare choices and reading time. In column (2), we thus report IV estimates using the policy change as instrument (see section 5.4.1). As in the linear regression, we find that both initial skills and quality of care are associated with increased child development. The IV estimates suggest, however, a larger role for both care modes in producing child human capital.

Table 3: Model parameters: child human capital and exhaustion effect

Parameter	Description	OLS (1)	IV (2)
δ_0	self-productivity	0.203***	0.105***
		(0.017)	(0.018)
δ_e	maternal care	0.154***	0.226***
		(0.026)	(0.020)
δ_d	non-maternal care	0.210***	0.264***
		(0.033)	(0.018)
$\gamma_{e,2}$	exhaustion effect	1.015	1.885***
		(0.022)	(0.205)
Exhaustion effect: IV First stage			
ITT($\ln(e)$)	parenting effort	0.0796***	
		(0.028)	
ITT($\ln(T_m)$)	maternal care	-0.0899***	
		(0.022)	

Note: This Table reports estimation results of model parameters for the child human capital production function (equation 9), the exhaustion-effect (equation 12), respectively. $\ln(h_1)$ is measured using the PPVT score, and is then standardized. The first-stage Sanderson-Windmeijer (SW) F-statistics for weak identification in the IV regression in column 2 are 36.34 for maternal care and 86.27 for non-maternal care. Bootstrapped standard errors (400 replications) are displayed in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
† $\gamma_{e,1}$ is re-scaled (multiplied) by 10,000 for comparability.

Exhaustion effect. In the last three rows, we display the estimation results of the convexity of the parenting-effort cost $\gamma_{e,2}$. Column (2) show the result of the IV-type estimator using the policy change discussed in Section 5.4.2. As derived earlier, the exhaustion-effect parameter is given by $\gamma_{e,2} = 1 - \frac{\text{ITT}(\ln(e))}{\text{ITT}(\ln(T_m))}$. ITT estimates of the policy's impact on log reading time and log maternal-care time are reported in the last two rows. We find that the reform led mothers to increase reading time by 0.08 log points and to reduce parenting time by 0.09 log points. These results suggest substitution between parenting time and effort, in line with CSW. They also imply a convexity in the cost of parenting effort (column 4) of $\gamma_{e,2} = 1.885$. Given that providing high-quality care is increasingly costly for parents, using childcare can provide some relief. Such reduction in the cost of parenting effort is potentially a key source of non-pecuniary gains for mothers.

Table 4: Preference parameters

Parameter	Description	Estimate	SE
γ_C	consumption	1	(.)
γ_ℓ	leisure	0.312***	(0.018)
γ_{T_m}	maternal care	1.872***	(0.178)
γ_{h_1}	child skills	2.499***	(0.080)
$\gamma_{e,1}$	cost of effort†	0.227***	(0.036)
$\gamma_{d,1}$	childcare use	1.643***	(0.156)
$\gamma_{d,2}$	coverage	1.615*	(0.879)

Note: This Table reports estimation results of preference parameters. Bootstrapped standard errors (400 replications) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
† $\gamma_{e,1}$ is re-scaled (multiplied) by 10,000.

Preferences. The final set of parameters is the preference vector γ. Table 4 shows the estimation results of the discrete-choice model (8). As expected, parents derive some utility from leisure time and time spent with the child. Moreover, our parameter estimates suggest that parents assign a significant value to their child's achievements but that parenting (effort) is costly.

Of particular interest from the perspective of the documented heterogeneity in policy impacts by local daycare supply, our estimates reveal that increased daycare coverage substantially reduces fixed costs on the childcare market. In the last two rows of Table 4, the parameter values indicate that in a hypothetical region with complete coverage (one space per preschool-age child) the entry cost would become negligible. These results again suggest a significant role for local daycare coverage in shaping parents' time-allocation choices. We now turn to a discussion of the model's fit before using our estimated model to compute the MVPF.

5.5.2 Model fit

To assess how accurate the model predictions are compared to the actual data, we use two approaches to test the validity of the model.

In-sample fit. First, in the first two columns of Table 5, we assess the model in-sample fit by comparing the time-allocation choices predicted by the model to observed parents' behavior in the pre-reform data. Using our parameter estimates of the three steps along with 200 draws of the extreme-value type-1 distribution for each household in the pre-reform Québec data, we create 200 datasets of predicted choices. We then compare key market-share summary statistics from the pre-reform data (column 1) to predicted statistics from our simulated dataset (column 2). In the first three rows, we examine the performance of our simulations at predicting extensive-margin choices. We find that the model is doing a decent job for maternal employment and the share of households reading weekly to the child, but slightly over-predicts childcare use. At the intensive margin (last two rows), we find that the model does not capture the small difference between the hours worked by the mother and childcare utilization that is observable in the data. This is likely due to the strong incentives in the model to take-up childcare when the mother works. When the mother works full-time, we assume the child must attend childcare at least part-time. Moreover, when the mother works part-time, she has to sacrifice hours of leisure if the household does not use childcare. Nevertheless, as Kaboski and Townsend (2011) argue, the model's ability to reproduce the reduced-form impacts of an intervention is arguably a stronger basis for evaluating a model's usefulness.

Table 5: Model fit

	In sample		Out-of-sample validation			
	Observed outcome (1)	Simulated outcome (2)	ITT estimates (3)	(4)	Simulated reform (5)	(6)
	Mean	Mean	Mean	CI	Mean	CI
Extensive margin						
Maternal employment	0.532	0.557	0.078	[0.013, 0.145]	0.072	[0.063, 0.080]
Childcare use	0.418	0.482	0.138	[0.074, 0.201]	0.110	[0.068, 0.146]
Reading weekly to child	0.748	0.680	0.044	[-0.023, 0.109]	0.004	[0.004, 0.004]
Intensive margin						
Mother's work hours	17.54	14.28	2.129	[-0.950, 5.237]	1.913	[1.684, 2.122]
Childcare hours	13.07	14.43	5.736	[3.629, 7.849]	3.482	[2.159, 4.632]

Note: This table displays the in-sample fit and results of our out-of-sample validation exercise, which compares the intent-to-treat estimates of the impact of the Québec reform (results in Table 1) with predictions from the policy simulation in the model. 95% confidence intervals (CI) on ITT estimates are computed using the Wild subcluster bootstrap of MacKinnon and Webb (2018) accounting for the small number of clusters. Confidence intervals on model predictions are computed using the simulation procedure of Krinsky and Robb (1986).

Out-of-sample validation. Thus, second, we perform an out-of-sample validation test by verifying whether the model predicts well the ITT estimates on maternal labor supply, childcare use, and time reading to the child. This validation exercise is similar in spirit to Chan and Liu (2018), who study a cash-for-care reform in Norway.[28]

We consider the three main aspects of the policy in our simulations: the price reduction in subsidized spaces, the local increase in supply, and the abolition of the refundable childcare credit. We map these features into the model by *(i)* reducing the hourly price of non-maternal care p to \$0.625 (assuming 8 hours of childcare per day), *(ii)* setting the credit rate τ_d to 0 for all households, and *(iii)* increasing the regional coverage rates $CovRate$ to their 2003 levels (second year of Cycle 5 data collection). Columns (3) to (6) of Table 5 present the results of this exercise by contrasting the predicted behavioral responses to this policy experiment in our simulation sample to the ITT estimates from Table 1.

For inference, since the model predictions are complex non-linear functions of preference parameters, we compute the standard errors using the simulation procedure of Krinsky and Robb (1986). We draw 1,000 parameter vectors from a multinomial normal distribution and predict behavioral responses for each draw. Confidence intervals for predictions are then

[28]This type of out-of-sample validation is also conducted by Todd and Wolpin (2006), Kaboski and Townsend (2011), and Chaparro et al. (2020), among others. See Todd and Wolpin (2023) for a review of empirical papers combining program evaluation with structural modeling.

obtained from quantiles of the simulated distribution of labor supply, childcare choices, and reading time.

We find that the model closely replicates the labor-supply response of mothers on both margins, and for childcare use. Indeed, our simulation of the policy predicts a 7.2 percentage points increase in maternal employment, which is very close to the reduced-form estimate of 7.8. Similarly, the model predicts an increase of 1.91 hours at the extensive margin, in line with the positive ITT estimate of 2.13 hours. For childcare use, the model also predicts a very similar impact on take-up, although the model slightly underpredicts childcare hours at the intensive margin. The model predicts no response of time reading to the child, in contrast with the positive impact found in the reduced-form analysis. Nevertheless, the good fit of mothers' labor supply and childcare take-up suggests that the model can rationalize key non-marginal responses.

6 Welfare analysis

In this section, we turn to the main contribution of the paper: we build on our reduced-form and our model estimates to compute the welfare impact of the policy. After briefly describing the welfare framework (Section 6.1), we construct the willingness to pay (Section 6.2), compute the policy's cost (Section 6.3), and present our MVPF estimates (Section 6.4). Lastly, Section 6.5 presents counterfactual estimates under alternative policy schemes.

6.1 Welfare framework

We build on the approach described in Hendren (2016) and Hendren and Sprung-Keyser (2020) where the social welfare impact of a policy change can be measured by its marginal value of public funds (MVPF), which is defined as the ratio between the policy's benefit to its beneficiaries (measured as their willingness to pay for that policy) and the policy's net cost to the government. That is:

$$\text{MVPF} = \frac{\text{Beneficiaries' Willingness to Pay}}{\text{Net Cost to Government}} \tag{13}$$

The net cost to the government is given by the difference between the direct government expenditure on the policy and fiscal externalities (i.e. indirect impacts on the government budget from changes in behavior). In our context, the main fiscal externality is the fiscal return from mothers' increased labor supply, which increases taxes collected and reduces transfers and benefits.

The higher the MVPF is, the larger the welfare gains to the beneficiaries per net dollar spent. A ratio higher than 1 indicates that the marginal benefit of the policy exceeds its marginal costs to taxpayers, suggesting that the policy is welfare-improving. An advantage of using this criterion is comparability across policy domains, in particular with targeted preschool programs studied in Hendren and Sprung-Keyser (2020). We now describe how we compute each component of the MVPF.

6.2 Willingness to pay

We start by the estimation of the numerator of the MVPF, a key contribution of this paper. To compute the WTP, it is crucial to distinguish two types of reforms: infinitesimal versus discrete policy changes. For sufficiently small policy changes, it can be shown that, under standard assumptions, the WTP boils down to the treatment effect on beneficiaries' after-tax earnings (Hendren, 2016). We illustrate this result below in the context of our model outlined in Section 5 and of our policy of interest. A more general exposition is provided in Appendix B.2.

6.2.1 Conceptual framework

Environment. Consider the model outlined in Section 5 in which the government chooses a childcare-provision policy characterized by a vector $\theta = (s, \pi)$, where s is a childcare subsidy (on the hourly price) and π is the childcare coverage rate. As in Hendren (2016), we assume that the labor and consumption-good markets are competitive so that the policy state θ has no impact on prices in those markets.[29] The decision problem thus depends on the policy state θ, which the mother takes as given.

We now abstract from the refundable childcare credit τ_d as it was replaced by a subsidy that is independent of income. Therefore, under the new regime, choices no longer depend on τ_d. The budget constraint thus now writes:

$$C + (p - s(\theta))\, T_d = wL + I \tag{14}$$

The agent takes prices (p and w), the coverage rate, the subsidy, non-labor income, the time budgets (T and T_c), non-maternal care quality (q), and the child's initial skills (h_0) as given and chooses consumption (C), labor hours (L), non-maternal care (T_d), and effort (e) to maximize utility. Solving the agent's maximization problem, we can write the indirect utility under θ as:

$$V(\pi(\theta), s(\theta), \omega) = \max_{C,L,T_d,e} U(C, L, T_d, e) \quad s.t.\ (2), (3), (5), (14)$$

where $\omega = (T, T_c, I, h_0, q, w, p)$ is a vector of exogenous variables. Suppose the government now implements a policy change, moving the policy state θ from the status-quo policy θ_0 to some new policy state θ_1. The agent's WTP for this policy change can be measured by the standard equivalent variation (E.V.). That is, the WTP is the variation in income under θ_0 that would make the agent indifferent between the status quo and the new policy state:

$$E.V. \;=\; \frac{V(\theta_1) - V(\theta_0)}{\lambda(\theta_0)} \tag{15}$$

where λ is the marginal utility of income.

[29] In reality, we might suspect a price response of daycare providers in the private (non-CPE) network to remain competitive. However, data we assembled from ministerial reports show that such pricing behavior is unlikely, at least over the time period considered in our analysis. As Appendix Figure A7 shows, average real daycare prices in the for-profit network remained relatively constant from 1994 to 1999.

WTP for a marginal policy change. Let us consider first a marginal (or infinitesimal) policy change in θ. The numerator in (15), the difference in indirect utilities, is then the total derivative of $V(\theta_0)$ with respect to θ. Among inframarginal agents, this yields:

$$\frac{dV(\theta_0)}{d\theta} = \underbrace{\gamma_{d,2}\frac{d\pi(\theta_0)}{d\theta}}_{\text{non-pecuniary gains}} + \underbrace{\lambda\, w\frac{dL(\theta_0)}{d\theta}}_{\text{pecuniary gains}} \tag{16}$$

Proof. See Appendix B.1.

Therefore, the numerator of the WTP is the sum of two terms: the *pecuniary* benefits, and the *non-pecuniary* gain stemming from the reduction in commuting costs. Under the assumption of a marginal policy change, the envelope theorem implies that, at the margin, behavioral responses do not have a direct effect on utility. It also implies that individuals who adjust their behavior as a response to the policy change must be indifferent between the two policy states. Thus, if the utility gains from the changes in coverage are negligible, which may be reasonable for marginal family reforms, the difference in utilities $(V(\theta_1) - V(\theta_0))$ boils down to the policy's impact on the beneficiaries' budget constraint. Thus, the treatment effect on beneficiaries' earnings is a sufficient statistic for the numerator of the MVPF (Hendren, 2016).

WTP for a non-marginal policy. Consider now a discrete (or non-infinitesimal) policy change. In this case, the previous result no longer holds since envelope conditions only apply to marginal reforms. This raises two concerns for the estimation of the WTP: compliers who adjust their labor supply may have direct utility gains, and non-pecuniary benefits are likely non-negligible. First, since agents make non-marginal changes in budgetary choices (i.e. they re-optimize their behavior), these no longer have a null direct impact on the difference in utilities $(V(\theta_1) - V(\theta_0))$. Second, focusing on earnings gains as an estimator of the WTP omits non-pecuniary benefits of the policy. This bias, in fact, also applies to marginal changes in θ if the first term in equation (16) is not null.

Social willingness-to-pay. To derive the society's WTP, one has to aggregate individual preferences taking into account preferences of the overall society. We focus on the case of a utilitarian planner who sets equal social weights to every agent.

6.2.2 Benchmark WTP

As a benchmark, we consider an estimator of the MVPF that assumes the Québec reform is marginal. By the envelope theorem, as shown above, the WTP is then simply the treatment effect on beneficiaries' earnings. Recall that the policy has no impact on fathers' labor supply so we assume there is no impact on non-labor income I.

To obtain this benchmark estimate, we combine our results on the short-term pecuniary impacts on mothers from Section 4.3 with other sources of pecuniary gains documented in the literature. Since only total net earnings gains matter under a utilitarian planner, we calculate the social WTP by simply multiplying the average effect by the number of mothers. This yields total short-term earnings gains of $2.469 billion.

To account for longer-term pecuniary gains, we also consider medium-run earnings gains of mothers. Lefebvre et al. (2009) estimate a positive effect on earnings of mothers of older children whose child was eligible when younger of \$1,995 over the period 1999-2004.[30] Multiplying these average earnings impacts by the number of mothers in two-parent families with children in those age ranges over that period, we obtain total earnings gains of about \$1.102 billion for mothers of older children.[31]

6.2.3 Structural WTP

We now account for re-optimization behavior and non-pecuniary gains using our structural model. This poses a key challenge in that we can no longer express the WTP as a single treatment effect parameter. Kleven (2021) shows that practitioners would need to estimate both "policy elasticities" and *changes* in elasticities along the policy path, which is arguably beyond empirical reach.

Hence, we use our estimated model to compute mothers' WTP. To do so, we simulate the reform in the model and estimate the WTP by computing mothers' equivalent variation as in Brink et al. (2007). The equivalent variation is the ratio of utility gains to the marginal utility of income (see equation 15). From the first-order condition for consumption, the marginal utility of income (λ) in our model is simply given by the inverse of the mother's total income net of childcare expenses. Denoting total income by $Y \equiv wL + I$, our structural estimator of the WTP thus writes:

$$\widehat{E.V.} = \frac{\widehat{V(\theta_1)} - \widehat{V(\theta_0)}}{\gamma_c / \left[\widehat{Y(\theta_0)} - \left(1 - \tau_d(\widehat{Y(\theta_0)})\right) p\widehat{T_d(\theta_0)}\right]} \tag{17}$$

where $\widehat{Y(\theta_0)}$ and $\widehat{T_d(\theta_0)}$ are the predicted income and childcare hours under the status quo respectively.

To measure this quantity in the model, we perform the following steps using our simulated sample. First, we obtain the indirect utility in the status quo for each synthetic mother by finding the alternative on the grid yielding the maximum utility. Status quo income net of childcare expenses is then given by the difference between the household's total income and total childcare expenses at that point on the grid. Second, we perform the policy simulation described in Section 5.5.2 where we simulate choices under the key reform parameters (the offer of a \$5/day spot, the increase in regional coverage, and the abolition of the refundable credit). The estimated indirect utility in the post-policy state $\widehat{V(\theta_1)}$ is then given by the new maximum utility on the grid. With all the estimated components in hand, we compute each synthetic mother's WTP using equation (17). Last, we take the average WTP over the 200 simulated duplicates of each mother observed in the NLSCY. To obtain the WTP for all Québec mothers of preschoolers, we use the sample weights provided by Statistics Canada. Because our simulation models the reform using 2003 coverage rates as the new policy state, to maintain comparability

[30] Like in our reduced-form analysis, we restrict dynamics in earnings impacts to that period to avoid capturing confounding effects of the Québec parental-leave reform of 2006, which had negative impacts on young mothers' earnings (Karademir et al., 2024).

[31] The data source for the number of mothers with children in given age ranges is the estimates from T1 Family Files of Statistics Canada (2023).

with the benchmark estimator that calculates average impacts over the roll-out of the daycare expansion, we sum the WTP over the pre-reform data only (two years).

Our estimates suggest the WTP is above 6 billion dollars, more than twice that in the benchmark. This result thus suggests that non-pecuniary gains are important in this context, which we further investigate through counterfactual simulations in Section 6.5.[32]

6.3 Net cost

We now consider the denominator of the MVPF, that is, the net cost of the policy for the government. As described above, this is given by the difference between the direct cost of the policy and fiscal externalities.

6.3.1 Direct cost

The upfront cost of the Québec reform is essentially comprised of the new subsidies allocated to the daycare market. These expenditures take various forms: start-up grants, recurring operating grants to daycare centres, special needs, and other subsidies. We first sum the total subsidies over the period covered in our empirical analysis, that is, over fiscal years 1997-1998 to 2002-2003. Figure A8 shows the evolution of total subsidies to daycare facilities along with the subsidy per space. Total subsidies increased from about \$470 million in fiscal year 1998-1999 to \$1.206 billion in 2002-2003. However, a share of these would have most likely been spent by the government absent the reform. Indeed, the government was spending nearly \$300 million annually in the two previous fiscal years. We assume similar subsidies would have been granted absent the policy change. We project these expenses assuming the same conservative growth rate observed from 1996 to 1997 (2.1%) and deduct these counterfactual subsidies from the observed grants. These calculations yield a total amount of new subsidies allocated of \$2.617 billion over our study period.[33]

6.3.2 Fiscal externality

We now calculate the government's return from behavioral changes. This includes fiscal benefits from increased maternal labor supply, which both expands the tax base – at both the extensive and intensive margins – and reduces tax credits and transfers due to higher household income decreasing eligibility.

Benchmark fiscal externality. We compute the fiscal return for the government using our estimates of the impact of the reform on earnings. Given that an increase in earnings has a differential effect on additional taxes paid and reduced benefits along the income distribution, we rely on our quantile regression analysis (see Section 4.3).

[32] Given that our model is one of the preschool period only, we cannot account for the medium-run on maternal labor supply in our structural estimator, as we did in the benchmark case. We are thus considering a lower bound for the WTP in the structural approach. This means that our MVPF estimates are conservative and might still underestimate the social return.

[33] A source of cost reduction is the potential savings from the abolition of the refundable childcare credits. Given that the Québec government allocated the same amount of public funds to this expenditure in 1996 and 2001, we abstract from potential savings from its abolition for families using a low-fee space and consider an upper bound on direct costs. A more detailed discussion is provided in Appendix A.3.

We perform the simulation in three steps. First, for each decile of the pre-reform mothers' income distribution in our sample, we compute the net fiscal position of the average mother in that decile, using the Canadian Tax and Credit Simulator (CtaCS) developed by Milligan (2019).[34] Then, we take the average earnings gain in a given decile and assign it to families in that decile. We then simulate the net additional taxes (net of transfers and benefits) paid by the mother under the post-reform (1998) tax parameters. This procedure yields an estimate of the fiscal externality of $856 million.

Structural fiscal externality. The marginal-policy assumption mostly affects the calculation of the WTP. However, to obtain an internally-consistent structural estimator of the MVPF, we also estimate the fiscal externality within our model. To do so, we calculate mothers' predicted income gains in our simulation of the policy. We then obtain their net fiscal positions using the CTaCS calculator. The total fiscal externality is then obtained by multiplying the simulated fiscal impact for each mother in the sample by their sample weight. We obtain an estimate of the fiscal externality that is slighlty larger in magnitude ($909 million) but comparable to our estimates using the benchmark causal effects. This similarity is reassuring and lends further support for the ability of our model to capture key policy impacts.

6.4 MVPF estimates

The MVPF is the ratio of the WTP to the net cost of the reform, which we computed in Sections 6.2 and 6.3, respectively. The net cost to society is the difference between the upfront expenditure and fiscal externalities. Table 6 displays the different components of our MVPF calculations, which we describe below.

Benchmark estimator. First, under the benchmark sufficient-statistic estimator, we obtain an estimate of the net cost of about $2,617M - $856M, which yields a net expense of $1.761 billion. Mothers' willingness-to-pay, captured by their earnings gains under the benchmark estimator, amounts to about $2.459 billion in after-tax income (substracting the fiscal externality above to the raw earnings gains). The benchmark estimator of the MVPF suggests mothers were willing to pay about $1.40 per net dollar spent on the reform. This is a small MVPF compared to targeted preschool programs studied in Hendren and Sprung-Keyser (2020).

Structural estimator. Second, our structural estimator suggests a similar, though slightly lower, fiscal externality from treated mothers of $909 million. However, the WTP differs substantially. Including non-pecuniary gains for mothers more than doubles the WTP. As a result, our estimate of the MVPF also more than doubles and reaches 3.56, bringing it considerably closer to the estimates reported by Hendren and Sprung-Keyser (2020) for targeted interventions.

We interpret these non-pecuniary utility gains as stemming from increased local availability of childcare, which reduces the costs associated with commuting time and the burden of finding

[34] The CTaCS is a comprehensive software that simulates the net fiscal position (at both provincial and federal levels of government) of an individual from a set of raw inputs (e.g. province, year, raw wage income, number of young children).

Table 6: Welfare estimates

MVPF components	Mean values	External source used
Direct cost	\$2,617M	Québec Treasury Board
Benchmark estimator		
Willingness-to-pay		
Mothers of preschoolers	\$2,213M	
Mothers of older children	\$1,102M	Lefebvre et al. (2009)
Taxes and reduced transfers	-\$856M	CTaCS
Fiscal externality		
Tax returns and reduced transfers	\$856M	CTaCS
MVPF	**1.40**	
Structural estimator		
Willingness-to-pay	\$6,078M	
Fiscal externality	\$909M	CTaCS
MVPF	**3.56**	

Notes: This table outlines the components of the MVPF under the benchmark and the structural estimators. The last column reports the external sources used for the policy's cost and other sources of fiscal externalities. The acronym CTaCS refers to the Canadian Tax and Credit Simulator of Milligan (2019).

a spot nearby. This interpretation is consistent with previous research showing that families highly value proximity to childcare, and that the implied willingness to pay for reduced travel time can be substantial. For instance, Bravo et al. (2022) estimate that the bulk of families' WTP for marginally closer childcare in Chile—282 USD out of a total of 343 USD—reflects parental time savings rather than children's human capital gains. Similarly, De Groote and Rho (2024) estimate that the costs associated with longer travel times to nurseries in Leuven, Belgium range from 0.83 to 1.9 EUR per minute and Borowsky (2019) estimates a WTP to avoid one mile of travel of \$2.65 in Minnesota. These findings support the view that commuting costs are a substantial component of childcare-related utility, in line with our own estimates.

Despite suggesting that the reform is welfare-improving, the sufficient-statistic estimator significantly underestimates the value of the Québec childcare reform for families—by more than half relative to our structural estimate.[35] This highlights a limitation of the sufficient-statistic approach when applied to non-marginal policies with important non-pecuniary benefits, a concern that potentially extends beyond childcare policy.[36]

[35] To account for uncertainty in our welfare estimates, we compute 95% confidence intervals for the benchmark estimator using the simulation method of Hendren and Sprung-Keyser (2020). That is, we take 10,000 draws from a multivariate normal distribution of the estimated policy impacts and obtain confidence intervals from quantiles of the simulated MVPF distribution. Unsurprisingly, we obtain large MVPF confidence bounds, ranging from 0.786 to $+\infty$, where an infinite MVPF means that the policy pays for itself. This is consistent with their findings, where most policies also exhibit confidence intervals too wide to be informative. Nonetheless, in over 75% of our simulations, our structural estimate exceeds the simulated sufficient-statistic one.

[36] See our survey in Appendix C for examples of such policies.

Robustness. We perform two sets of robustness checks on our main welfare estimates. First, we consider long-run impacts on children's behavior. We recall that we did not find evidence of pecuniary gains for children as they age (Section 4.2). However, previous literature suggests other long-term consequences. In particular, Baker et al. (2019) document that the policy coincides with an increase in youth crime at ages 12-20 among exposed cohorts as they aged.[37] As a robustness check, we monetize these additional societal costs in Appendix A.7 using estimates of costs of juvenile crime of Cohen (2020). We obtain a WTP to avoid these transgressions (victim costs and reduced productivity of the offender) of $20.16 million (in 1997 dollars) and a fiscal externality that is similar in magnitude. Thus, these impacts on children are somewhat negligible compared to mothers' earnings gains and direct costs which amount to billions of dollars. As a result, the benchmark MVPF only slightly decreases to 1.37. Appendix A.7 and Table A8 report further details on these calculations.

Second, we compare our MVPF estimates to other standard metrics, namely the benefit-cost ratio (BCR) and the net social benefit (NSB), in Appendix A.8. This notably allows us to verify the robustness of our results to adjusting for the deadweight loss of raising public funds. While the MVPF, by construction, does not apply this adjustment (see Hendren and Sprung-Keyser, 2022, and Appendix A.8 for discussions), we find that using the BCR or the NSB confirms our main conclusions that benefits exceed costs under the structural estimator.

Limitations. Lastly, we acknowledge that our analysis is subject to a few important limitations. First, while structural modeling allows us to relax the marginal-policy assumption of sufficient-statistics estimators, this approach comes with assumptions on model primitives. As in any structural analysis, our results rely on the functional forms chosen, identification from choice behavior, and assumptions about preference stability. However, the out-of-sample validation of our model—showing that we match key reduced-form impacts of the reform—bolsters our confidence in the validity of the model and our conclusions.

Second, our analysis focuses on short-term outcomes for mothers and does not capture potential longer-run effects on maternal labor supply. However, to the extent that childcare provision generates dynamic longer-run benefits, such as career progression and increased lifetime earnings, our estimates likely represent a lower bound on the long-term welfare gains for mothers. We also limit our analysis to mothers, excluding potential impacts on children, fathers, or intra-household bargaining dynamics. While this is clearly a simplification, we find no significant average effects of the reform on fathers, and we can safely rule out large negative effects on children as they age. Therefore, we believe that our focus on short-term impacts on mothers provides a conservative estimate of the overall welfare gains from universal childcare.

Third, our data lack validated measures of childcare quality. While we observe proxies for parental perceptions of quality, we do not have access to expert-validated quality scales such as those used in Griffen (2019), Chaparro et al. (2020), Bodéré (2023), or Berlinski et al. (2024). Therefore, we are unable to analyze counterfactual scenarios varying quality standards as some of these studies do for other programs. In such scenarios, higher quality is typically associated with improved long-term outcomes for children, but also with increased program costs. For

[37] To be sure, prevalent youth crimes are rather minor transgressions such as thefts of small amounts, mischief, breaking and entering, failure to appear in court, and cannabis possession (Baker et al., 2019).

example, the iconic Carolina Abecedarian Project studied in García et al. (2020) cost \$18,514 per participant (in 2014 USD), yet yielded significantly larger lifetime benefits to targeted children. Nevertheless, we do not expect that higher quality would have further increased childcare demand in our setting, given that demand was already high and that parents typically struggle to recognize high-quality childcare (Herbst et al., 2020; Gordon et al., 2021).

6.5 Policy counterfactuals

In this Section, we use our model to assess *(i)* the contribution of each policy feature to the estimated welfare gains, and *(ii)* whether the government could have obtained higher welfare gains under alternative policy schemes.

Mechanism. First, we ask which feature of the policy is responsible for the bulk of the welfare gain or, in other words, for which policy feature mothers are relatively more willing to pay for. We simulate counterfactual scenarios in which we separately remove each feature of the policy and compute the associated WTP. We simulate, in turn, counterfactual policies under which families face *(i)* the higher pre-reform prices and *(ii)* the lower pre-reform coverage rates.

Our results, displayed in Table 7 suggest that most of the welfare gains are due to increased coverage. In the second counterfactual, the WTP for increasing coverage rates to their 2003 level without decreasing the price is 84% that of the actual reform (\$5,120M). However, the WTP for the price reduction only—our first counterfactual—is very small (only \$362M) compared to that of the actual policy, suggesting that the decrease in price is not the main driver of welfare gains. We stress that this result is not simply mechanical. In our model, an increase in coverage only reduces entry costs in the childcare market and thus does not increase the number of spaces available *per se*. Therefore, the WTP for coverage reflects the value of providers' proximity. Our results thus suggest that increasing childcare availability is key for the effectiveness of universal preschool policies.[38]

Table 7: Counterfactual willingness-to-pay

Policy scenario	WTP	Share of structural WTP
Actual reform		
Benchmark	\$2,344M	38.6%
Structural	\$6,078M	100%
Counterfactuals		
No price decrease	\$5,120M	84.2%
No coverage increase	\$362M	6%

Notes: This table outlines WTP estimates under the actual reform and counterfactual scenarios, namely cases maintaining pre-reform (*i*) higher prices, and (*ii*) lower coverage rates.

[38]Those results are in line with De Groote and Rho (2024), who find large welfare gains of increasing daycare capacity on a centralized Belgian platform. In particular, they show that even a small increase in daycare capacity is sufficient to compensate advantaged families for their welfare loss under affirmative-action policies.

Alternative policies. Second, we compare the MVPF of the adopted reform to counterfactual MVPFs of alternative policy schemes under difference price-coverage combinations.

To this end, we simulate behavioral responses under multiple price-coverage pairs. For simplicity, we assume a uniform coverage rate throughout the province. For each pair, we proceed in three steps as follows. In a first step, we obtain mothers' WTP by calculating the counterfactual equivalent variation (17) using our synthetic datasets. In a second step, we compute the counterfactual fiscal externality using the CTaCS calculator following the same approach as for the actual reform.

In the third step, we compute the counterfactual direct costs. To keep the computations tractable, we make the following simplifying assumptions: (i) government subsidies vary dollar-for-dollar with the expenses made by families, and (ii) daycare centres operate over the maximum number of days (260 days).[39] We nevertheless take into account the fact that, in counterfactual scenarios with high coverage rates, not all spaces are filled, which means that the government alone pays for unused spaces.[40]

The results are reported in Figure 6, which shows how the simulated MVPF, WTP and net cost—where those of the actual reform are normalized to 1—vary with price and coverage. We find a striking pattern: social welfare gains are generally increasing in daycare coverage but also with the fee charged to families. Panel (b) shows that this is because mothers' WTP is less sensitive to the daycare price than to coverage. While this result might be context-dependent, it is consistent with evidence from other daycare-subsidy policies where shortages were present.[41]

We stress, however, that our empirical model is a partial-equilibrium framework and abstracts from general-equilibrium effects on other childless workers. Using the Canadian Labour Force Surveys, we do not find differential trends in hourly wages between Québec and other provinces among childless women and men, suggesting a limited role for such general-equilibrium wage effects (see Appendix Figure A9). Nevertheless, in counterfactual simulations with large coverage-rate increases, the model predicts such large impacts on maternal labor supply that abstracting from general-equilibrium effects would likely be unrealistic. Large increases in labor supply would plausibly exert downward pressure on wages in the overall economy.[42] Increased labor demand in the childcare sector would also put upward pressure on

[39] This is consistent with the typical childcare contracts in Québec over that period. To be eligible to the universal subsidy, families were required to enrol their child full-time for a maximum of 260 days per year, but families would typically sign yearly contracts to keep their space. The fees had to be paid even if the child was absent from daycare. Note, nevertheless, that the for-profit market was not much more flexible, providers preferring to offer only full-time spaces because they were easier to manage (Haeck et al., 2018).

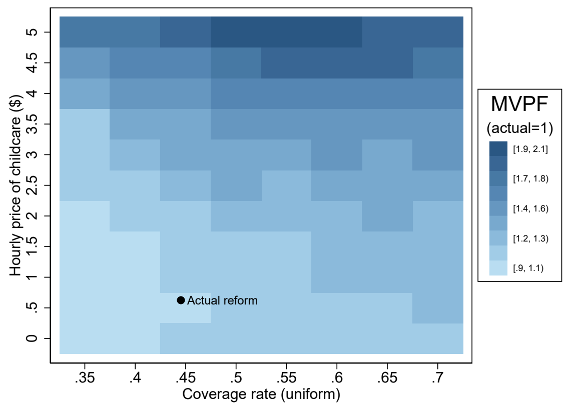
[40] Counterfactual government subsidies, denoted by G', are then given by the following expression:

$$\begin{aligned} G' &= (\text{Total costs})' - (\text{Parents' payments})' \\ &= (G + \$5 \times \text{spaces} \times 260) \times \frac{(CovRate)'}{CovRate} - p' \times 260 \times min\left\{(\text{childcare use})', (CovRate)'\right\} \times pop \end{aligned}$$

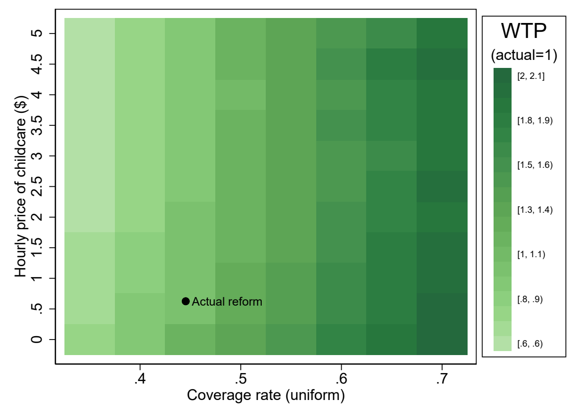
where *pop* is the population of preschool-age children in the post-policy period and G is the total government subsidy under the actual reform. Variables with a $'$ sign refer to the values of these variables under counterfactual scenarios.

[41] Yamaguchi et al. (2018) document that, in the case of the Japanese national childcare expansion, pecuniary gains were increasingly larger as the program expanded.

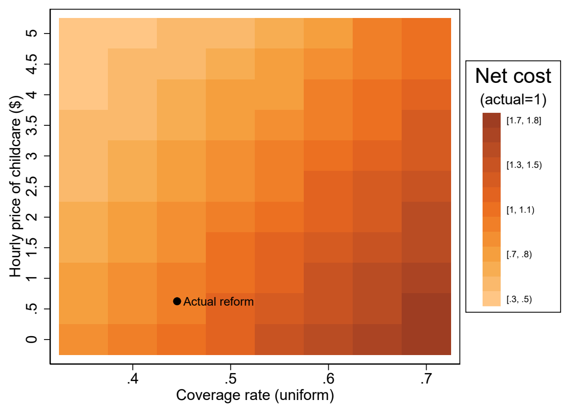
[42] Prices in daycare are not a relevant margin for general-equilibrium effects in our setting. We focus on the period during which the private market could not expand, and prices were fixed in the public network. We also



(a) MVPF



(b) WTP



(c) Net cost

Figure 6: MVPF, WTP, and net cost under counterfactual price-coverage combinations

Note: This figure plots the simulated counterfactual willingness-to-pay (WTP), net cost, and Marginal Value of Public Funds (MVPF) under different daycare price and coverage rate combinations. Estimates for the actual reform are normalized to 1. The hourly price of $0.625 for the actual reform assumes eight hours of care per day. Darker colors represent higher values.

wages in that sector, thereby raising the costs of public provision. For these reasons, we interpret results from large counterfactual expansions with caution and do not evaluate scenarios with coverage rates above 0.7.

Overall, our findings suggest that increasing childcare availability is key for the effectiveness of universal preschool policies. Mothers value availability the most, and the government could have achieved larger welfare gains by channeling more resources towards opening new spots rather than to lowering prices. For instance, compared to the actual reform, counterfactual estimates suggest the MVPF for a reform that doubles the price charged to families and increases the coverage rate by 5 percentage points would be 13% larger.

7 Conclusion

Often inspired by impressive results of childcare programs targeted at disadvantaged families, many countries have proposed offering affordable childcare alternatives to all. Yet, available

do not observe significant changes in prices among private providers from 1994 to 1999 (see Figure A7).

evidence on the effectiveness of universal reforms has suggested much smaller social returns. In this paper, we provide a comprehensive welfare analysis of a major universal preschool reform in Québec and show that benefits have been underestimated. Moving beyond the fiscal impact of such policy change, we estimate the value of the reform for mothers using a structural model of childcare demand. Our estimates indicate that the benefit-to-net-cost ratio is more than twice as large when non-pecuniary gains are considered via this approach.

Our results suggest three lessons for empirical welfare analysis of universal preschool reforms. First, these policies can yield substantial welfare gains, in particular in the form of non-pecuniary benefits for mothers. Thus, focusing on fiscal externalities to the government substantially underestimates the social return. Second, the study highlights the limitations of sufficient-statistic methods in welfare analysis, often implicitly used in benefit-cost analyses of large reforms. We show that, when applied to non-marginal preschool policies, this approach might omit key welfare gains which matter empirically. Third, it underscores the importance of improving local availability for the success of universal programs. While affordability of childcare has received much attention from scholars and policymakers, our results suggest that family policy supporting supply in childcare markets is at least as important.

We conclude by highlighting some research avenues. One key message from this paper is that availability of childcare options is crucial for families. The literature on the determinants of daycare supply is still in its infancy and more work is needed to understand trade-offs providers face in the daycare market. In particular, a better understanding of which type of family policy can ensure that providers offer high-quality services is crucial. Borowsky et al. (2022), Bodéré (2023), and Berlinski et al. (2024) make advances on this front by modeling the supply side of the US childcare market. However, it is still unclear whether results from a market-oriented context like the US also apply to other settings with a long history of public provision such as Europe. Our study also sheds light on the empirical limitations of sufficient-statistic metrics for welfare analysis. Studying non-marginal policies, however, comes at the "cost" of structural assumptions on the economic problem and perhaps realism. Developing intermediate approaches that preserve the benefits of causal inference methods appears as a fruitful area for future research.

References

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Andresen, M. E. and Havnes, T. (2019). Child care, parental labor supply and tax revenue. *Labour Economics*, 61:101762.

Apps, P., Kabátek, J., Rees, R., and van Soest, A. (2016). Labor supply heterogeneity and demand for child care of mothers with young children. *Empirical Economics*, 51(4):1641–1677.

Bailey, M. J., Sun, S., and Timpe, B. (2021). Prep school for poor kids: The long-run impacts of head start on human capital and economic self-sufficiency. *American Economic Review*, 111(12):3963–4001.

Baker, M. (2011). Innis lecture: Universal early childhood interventions: what is the evidence base? *Canadian Journal of Economics/Revue canadienne d'économique*, 44(4):1069–1105.

Baker, M., Gruber, J., and Milligan, K. (2005). Universal childcare, maternal labor supply, and family well-being. Working Paper 11832, National Bureau of Economic Research.

Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745.

Baker, M., Gruber, J., and Milligan, K. (2019). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy*, 11(3):1–26.

Becker, G. S. (1965). A theory of the allocation of time. *The economic journal*, 75(299):493–517.

Becker, G. S. (1991). *A treatise on the family: Enlarged edition.* Harvard university press.

Bergstrom, K., Dodds, W., and Rios, J. (2025). Welfare analysis of changing notches: Evidence from bolsa família. *American Economic Journal: Economic Policy*, 17(2):122–61.

Berlinski, S., Ferreyra, M. M., Flabbi, L., and Martin, J. D. (2024). Childcare markets, parental labor supply, and child development. *Journal of Political Economy*, 132(6):2113–2177.

Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. *International Economic Review*, 49(4):1173–1209.

Bernal, R. and Keane, M. P. (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics*, 156(1):164–189.

Blau, D. M. and Hagy, A. P. (1998). The demand for quality in child care. *Journal of Political Economy*, 106(1):104–146.

Bodéré, P. (2023). Dynamic spatial competition in early education: an equilibrium analysis of the preschool market in pennsylvania. Mimeo.

Borowsky, J. (2019). Who benefits from child care ratings? evidence from minnesota's parentaware program. Technical report, Department of Economics, University of Minnesota, Minneapolis and Saint Paul, MN.

Borowsky, J., Brown, J. H., Davis, E. E., Gibbs, C., Herbst, C. M., Sojourner, A., Tekin, E., and Wiswall, M. J. (2022). An equilibrium model of the impact of increased public investment in early childhood education. Working Paper 30140, National Bureau of Economic Research.

Bravo, D., Hojman, A., and Rodríguez, J. (2022). Welfare effects of childcare policies. Working paper.

Brink, A., Nordblom, K., and Wahlberg, R. (2007). Maximum fee versus child benefit: a welfare analysis of swedish child-care fee reform. *International Tax and Public Finance*, 14:457–480.

Buera, F. J., Kaboski, J. P., and Townsend, R. M. (2023). From micro to macro development. *Journal of Economic Literature*, 61(2):471–503.

Caucutt, E. M. and Lochner, L. (2020). Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy*, 128(3):1065–1147.

Chan, M. K. and Liu, K. (2018). Life-cycle and intergenerational effects of child care reforms. *Quantitative Economics*, 9(2):659–706.

Chaparro, J., Sojourner, A., and Wiswall, M. J. (2020). Early childhood care and cognitive development. Working Paper 26813, National Bureau of Economic Research.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(1):451–488.

Cleveland, G., Forer, B., Hyatt, D., Japel, C., and Krashinsky, M. (2008). New evidence about child care in canada: Use patterns, affordability and quality. 14(12).

Cleveland, G. and Krashinsky, M. (2009). The nonprofit advantage: Producing quality in thick and thin child care markets. *Journal of Policy Analysis and Management*, 28(3):440–462.

Cohen, M. A. (2020). *The costs of crime and justice*. Routledge.

Cornelissen, T., Dustmann, C., Raute, A., and Schönberg, U. (2018). Who benefits from universal child care? estimating marginal returns to early child care attendance. *Journal of Political Economy*, 126(6):2356–2409.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American economic review*, 97(2):31–47.

Daruich, D. (2022). The macroeconomic consequences of early childhood development policies. Working paper.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

De Groote, O. and Rho, M. (2024). Matching disadvantaged children to day care: Evidence from a centralized platform. TSE Working Paper 24-1600, Toulouse School of Economics.

Del Boca, D., Flinn, C., and Wiswall, M. (2014). Household choices and child development. *Review of Economic Studies*, 81(1):137–185.

Ding, W., Kottelenberg, M. J., and Lehrer, S. F. (2021). Not quite universal childcare: Evidence on how the demand and supply side initially respond to a shortage of subsidized spaces in childcare. Technical report, Working Paper.

Duncan, G., Kalil, A., Mogstad, M., and Rege, M. (2023). Investing in early childhood development in preschool and at home. In *Handbook of the Economics of Education, Volume 6*, pages 235–297. Elsevier.

Felfe, C. and Lalive, R. (2018). Does early child care affect children's development? *Journal of Public Economics*, 159:33–53.

Finkelstein, A. and Hendren, N. (2020). Welfare analysis meets causal inference. *Journal of Economic Perspectives*, 34(4):146–67.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.

Fitzpatrick, M. (2010). Preschoolers enrolled and mothers at work? the effects of universal prekindergarten. *Journal of Labor Economics*, 28(1):51–85.

Fort, M., Ichino, A., and Zanella, G. (2020). Cognitive and noncognitive costs of day care at age 0–2 for children in advantaged families. *The Journal of Political Economy*, 128(1):158–205.

Fortin, P., Godbout, L., and St-Cerny, S. (2013). L'impact des services de garde à contribution réduite du québec sur le taux d'activité féminin, le revenu intérieur et les budgets gouvernementaux. *Revue Interventions économiques. Papers in Political Economy*, 47.

García, J. L., Heckman, J. J., Leaf, D. E., and Prados, M. J. (2020). Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy*, 128(7):2502–2541.

García, J. L., Heckman, J. J., and Ronda, V. (2023). The lasting effects of early-childhood education on promoting the skills and social mobility of disadvantaged african americans and their children. *Journal of Political Economy*, 131(6):1477–1506.

Gelbach, J. B. (2002). Public schooling for young children and maternal labor supply. *American Economic Review*, 92(1):307–322.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American economic review*, 96(2):1–21.

Gordon, J. A., Herbst, C. M., and Tekin, E. (2021). Who's minding the kids? experimental evidence on the demand for child care quality. *Economics of Education Review*, 80:102076.

Griffen, A. S. (2019). Evaluating the effects of childcare policies on children's cognitive development and maternal labor supply. *Journal of Human Resources*, 54(3):604–655.

Guner, N., Kaygusuz, R., and Ventura, G. (2020). Child-related transfers, household labour supply, and welfare. *The Review of Economic Studies*, 87(5):2290–2321.

Haeck, C., Lebihan, L., Lefebvre, P., and Merrigan, P. (2022). Childcare, parental behavior and well-being during childhood. *The BE Journal of Economic Analysis & Policy*, 22(2):311–360.

Haeck, C., Lebihan, L., and Merrigan, P. (2018). Universal child care and long-term effects on child well-being: Evidence from canada. *Journal of Human Capital*, 12(1):38–98.

Haeck, C., Lefebvre, P., and Merrigan, P. (2015). Canadian evidence on ten years of universal preschool policies: The good and the bad. *Labour Economics*, 36:137–157.

Haeck, C., Lefebvre, P., and Merrigan, P. (2016). La politique des services de garde à contribution réduite (sgcr) : retombées probantes, faiblesses et pistes d'améliorations. Unpublished manuscript,

Groupe de recherche sur le capital humain.
Havnes, T. and Mogstad, M. (2011a). Money for nothing? universal child care and maternal employment. *Journal of Public Economics*, 95(11):1455–1465.
Havnes, T. and Mogstad, M. (2011b). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy*, 3(2):97–129.
Havnes, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, 127:100–114.
Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–86.
Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the highscope perry preschool program. *Journal of Public Economics*, 94(1):114–128.
Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy*, 30(1):51–89.
Hendren, N. and Sprung-Keyser, B. (2020). A Unified Welfare Analysis of Government Policies*. *The Quarterly Journal of Economics*, 135(3):1209–1318.
Hendren, N. and Sprung-Keyser, B. (2022). The case for using the mvpf in empirical welfare analysis. Working Paper 30029, National Bureau of Economic Research.
Hendren, N., Sprung-Keyser, B., Stuart, E., et al. (2025). Policy impacts library. Last accessed: 2023-01-16.
Herbst, C. M. (2017). Universal Child Care, Maternal Employment, and Children's Long-Run Outcomes: Evidence from the US Lanham Act of 1940. *Journal of Labor Economics*, 35(2):519–564.
Herbst, C. M., Desouza, K. C., Al-Ashri, S., Kandala, S. S., Khullar, M., and Bajaj, V. (2020). What do parents value in a child care provider? evidence from yelp consumer reviews. *Early Childhood Research Quarterly*, 51:288–306.
Hojman, A. and Lopez Boo, F. (2022). Public childcare benefits children and mothers: Evidence from a nationwide experiment in a developing country. *Journal of Public Economics*, 212:104686.
Hotz, V. J. and Wiswall, M. (2019). Child care and child care policy: existing policies, their effects, and reforms. *The ANNALS of the American Academy of Political and Social Science*, 686(1):310–338.
Institut de la Statistique du Québec (2023). Statistiques par région administrative. Last accessed 1 November 2023.
Japel, C., Tremblay, R. E., and Côté, S. (2005). Quality counts! assessing the quality of daycare services based on the quebec longitudinal study of child development. *Choices: Investing in Our Children*, 11(5).
Kaboski, J. P. and Townsend, R. M. (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica*, 79(5):1357–1406.
Kang, Z. Y. and Vasserman, S. (2024). Robustness measures for welfare analysis. Technical report, National Bureau of Economic Research.
Karademir, S., Laliberté, J.-W., and Staubli, S. (2024). The multigenerational impact of children and childcare policies. *Journal of Labor Economics*. forthcoming.
Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2024). Do family policies reduce gender inequality? evidence from 60 years of policy experimentation. *American Economic Journal: Economic Policy*, 16(2):110–149.
Kleven, H. J. (2021). Sufficient statistics revisited. *Annual Review of Economics*, 13(1):515–538.
Kline, P. and Walters, C. R. (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start*. *The Quarterly Journal of Economics*, 131(4):1795–1848.
Kottelenberg, M. J. and Lehrer, S. F. (2017). Targeted or universal coverage? assessing heterogeneity in the effects of universal child care. *Journal of Labor Economics*, 35(3):609–653.
Krinsky, I. and Robb, A. L. (1986). On approximating the statistical properties of elasticities. *The review of economics and statistics*, 68(4):715–719.
Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from canada. *Journal of Labor Economics*, 26(3):519–548.
Lefebvre, P., Merrigan, P., and Verstraete, M. (2009). Dynamic labour supply effects of childcare subsidies: Evidence from a canadian natural experiment on low-fee universal child care. *Labour Economics*, 16(5):490–502.
List, J. A., Suskind, D., and Supplee, L. H. (2021). *The Scale-Up Effect in Early Childhood and Public Policy: Why Interventions Lose Impact at Scale and What We Can Do About It*. Routledge.
MacKinnon, J. G. and Webb, M. D. (2018). The wild bootstrap for few (treated) clusters. *The Econometrics Journal*, 21(2):114–135.

McFadden, D. L. (1974). Conditional logit analysis of qualitative choice behavior. In P, Z., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press.
Milligan, K. (2019). Canadian tax and credit simulator. Database, software and documentation, Version 2019-1.
Ministère de la Famille (2013–2019). Rapport annuel. Technical report, Gouvernement du Québec.
Ministère de la Famille, des Aînés et de la Condition Féminine (2005–2006). Rapport annuel de gestion. Technical report, Gouvernement du Québec.
Ministère de la Famille et de l'Enfance (1997–2003). Rapport annuel de gestion. Technical report, Gouvernement du Québec.
Ministère de la Famille et de l'Enfance (1998–2000). Situation des centres de la petite enfance et des garderies au québec. Technical report, Gouvernement du Québec.
Ministère de la Famille et de l'Enfance (2001). Situation des centres de la petite enfance et des garderies au québec. Technical report, Gouvernement du Québec.
Ministère de la Famille et des Aînés (2007–2012). Rapport annuel de gestion. Technical report, Gouvernement du Québec.
Ministère de l'Emploi, de la Solidarité sociale et de la Famille (2004). Rapport annuel de gestion 2003-2004. Technical report, Gouvernement du Québec.
Molnár, T. L. (2023). Costs of daycare, complementarities, and heterogeneous productivity of parenting time in child skill formation. Working paper.
Office des services de garde à l'enfance (1994–1996). Estimation régionale des besoins en services de garde régis par la loi. Technical report, Gouvernement du Québec.
Office des services de garde à l'enfance (1994–1997c). Situation des centres de la petite enfance et des garderies au québec. Technical report, Gouvernement du Québec.
Office des services de garde à l'enfance (1997a). Estimation régionale des besoins en services de garde régis par la loi. Technical report, Gouvernement du Québec.
Office des services de garde à l'enfance (1997b). Situation des centres de la petite enfance et des garderies au québec. Technical report, Gouvernement du Québec.
Office des services de garde à l'enfance et Secrétariat à la famille (1997). Rapport annuel 1996-1997. Technical report, Gouvernement du Québec.
Patnaik, A. (2019). Reserving time for daddy: The consequences of fathers' quotas. *Journal of Labor Economics*, 37(4):1009–1059.
Penn, H. and Lloyd, E. (2012). *Childcare Markets: Can They Deliver an Equitable Service?* Bristol University Press.
Ribar, D. C. (1995). A structural model of child care and the labor supply of married women. *Journal of labor Economics*, 13(3):558–597.
Seward, B., Dhuey, E., and Pan, A. (2023). The big short: Expansion of early childhood education in post-pandemic canada. *Canadian Public Policy*, 49(3):306–329.
Statistics Canada (1999). Census of population 1996. Data file.
Statistics Canada (2010). National longitudinal survey of children and youth. Data file.
Statistics Canada (2018). Census of population 2016. Data file.
Statistics Canada (2022). Census of population 2021. Data file.
Statistics Canada (2023). Table 39-10-0041-01 census families with children by age of children and children by age groups. Last accessed 6 September 2023.
Todd, P. E. and Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American economic review*, 96(5):1384–1417.
Todd, P. E. and Wolpin, K. I. (2023). The best of both worlds: Combining randomized controlled trials with structural modeling. *Journal of Economic Literature*, 61(1):41–85.
Yamaguchi, S., Asai, Y., and Kambayashi, R. (2018). Effects of subsidized childcare on mothers' labor supply under a rationing mechanism. *Labour Economics*, 55:1–17.

A Additional results and background information

A.1 Quality of care in childcare facilities

Under strong public pressure to open more spots at a reduced fee, the Québec government maintained minimal educational standards for daycare workers to facilitate entry into the profession at the implementation of the reform. A report on childcare quality in regulated settings from the *Institut de la Statistique du Québec* (Québec Statistical Institute) in 2004 emphasized the need for improving quality of care in those institutions (see Drouin et al., 2004). Moreover, an audit study conducted by Japel et al. (2005) between 2000 and 2003 revealed that the majority of childcare settings (61%) only met the basic criteria (ensuring the children's health and safety), and that their educational component was minimal. Almost one-eighth of them failed to meet the minimum standards.

However, Japel et al. (2005) also found that CPEs, on average, outperformed all other childcare settings on the vast majority of the criteria they considered for quality. For example, 26.5% of unregulated daycares (home-based or for-profit) were rated as inadequate in terms of quality, but only 6% of CPEs were rated as such. In the same vein, only 12.5% of unregulated daycares provided more than the "minimal" quality, while 33% of CPEs were deemed to provide a good service. Therefore, this evidence suggests that quality issues were actually more important in the private childcare market.[43] One part of the solution to improve average quality, some observers argued, was thus to increase the number of reduced-fee regulated spaces in CPEs.

In response to these quality issues, the provincial government, in addition to increasing the quantity of subsidized spaces, also gradually implemented some quality changes. In 2000, as documented by Molnár (2023), the educational requirements and wages of staff in regulated facilities were substantially increased over a four-year period. The average wage of child care workers was raised by 38 to 40 percent over this time span. The staff-to-child ratios remained unchanged (except for four and five year olds whose ratio increased by 25%) despite the increase in maximum capacity (Baker et al., 2005). Qualification requirements for the staff in centre-based CPEs were raised, and they were then extended to all centre-based care in 2006. In addition to political will, the increase in parents' involvement (in the board of directors, for example) is also an important factor that led to these quality changes.

A.2 Brief summary of microdata sources

The NLSCY studies the development and well-being of Canadian children. Children are followed bi-annually from birth to early adulthood and the information collected includes a range of indicators of socio-emotional, cognitive, and behavioral development. The survey series began in 1994 with an initial sample of children under the age of 12, which was followed for 14 years (at two-year intervals) through 2008. In each wave of the survey, a new cohort of children under the age of two was added to focus on early childhood development. On top of including detailed information on child development outcomes, the survey also includes a questionnaire

[43] One potential reason for those quality differences is the greater generosity of infrastructure subsidies given by the government to daycare providers of subsidized spaces than to those operating privately.

given to the parent who is deemed the most knowledgeable about the child. In that section, the parent provides information on family functioning, parental support, labor supply, and the time allocation of care among various options (daycare center, at home by relative, at home by non-relative, and so on). Weekly expenditures on care are also reported by the parent in the last two waves and we use these data to infer the price of private-market care.

The Canadian Census of Population is conducted every five years since 1981 and collects information on all household members from a 20% to 25% sample representative of the Canadian population. It covers a variety of topics, of which we notably extract the province of birth, demographic characteristics as well as educational attainment. We use the 2016 and 2021 waves of the Census for the empirical analysis of children's long-run outcomes.

The Longitudinal Administrative Databank (LAD) is a longitudinal file comprised of a 20% sample of annual tax declarations (the T1 Family Files) of Canadians who have a social insurance number. It contains detailed information on individuals' sources of income, tax credits, and taxes paid along with some demographics such as household composition. We use the LAD data covering tax years 2010 to 2022 for our long-run analysis of children's earnings.

The Canadian Labor Force Surveys (LFS) are annual surveys of the working-age population – excluding some specific categories of Canadian households (individuals in Aboriginal reserves, members of the Canadian Forces, and the institutionalized) – that include labor-market and basic demographic variables. The data contain hourly wages for employed individuals starting in 1997. We use the summary Table 37-10-0130-01 of Statistics Canada using the LFS for plots of long-run trends in education across Canadian provinces reported in Figure A10.

A.3 Abolition of the refundable childcare credit

To lower the impact of subsidized daycare on public finances, the Québec government made simultaneous changes to other family allowances. In particular, for families obtaining a low-fee space, the refundable childcare credit available before the policy was abolished. To obtain an estimate of the savings generated by this fiscal policy change, we compare government expenses in this expenditure item before and after the reform. Total childcare credits allocated to families are retrieved from the Québec Ministry of Finance's annual budget. In 1996, the fiscal year just before the reform, the credit cost $192 million. In 2001, the same amount was allocated to this program ($191 million). Given the sharp decrease in the number of unsubsidized spaces (who became subsidized), this pattern is surprising. Perhaps some families, who were not claiming the credits before the reform, suddenly started doing so. We are thus reluctant to attempt to impute what would have been spent by the government absent the Québec reform. Those savings would most likely be small in any case compared to the size of the subsidies (Fortin et al., 2013). We thus prefer to consider an upper bound on direct costs and abstract from potential savings from this source in the analysis.

A.4 Alternative specification for regional heterogeneity

In this Appendix, we consider an alternative specification to investigate heterogeneity in policy impacts by local daycare supply. We estimate the following model:

$$Y_{iprt} = \alpha + \beta_1 \text{Elig}_{pt} + \beta_2 \text{Elig}_{pt} \times \text{LowExp}_r + \gamma_p + \gamma_r + \gamma_t + \sigma W_{rt} + \delta X_{ipt} + \varepsilon_{iprt} \quad (18)$$

where r indicates the administrative region of residence (within Québec only). LowExp_r equals one if region r is in the bottom tercile of the distribution of daycare expansion over the period. This variable aims at capturing regions in which daycare expanded little. The expansion level is defined as the difference between region r's daycare coverage rate in 2003 to its 1997 level. γ_r is a vector of region (within Québec only) fixed effects. Lastly, W_{rt} is a vector of regional-level control variables associated with childcare demand (shares of medium- and high-educated mothers and the number of preschoolers in the region).

Results. The main coefficients of interest are β_1 and β_2, which capture the differential effects of the policy by local daycare availability. Tables A9 and A10 report point estimates from this specification, along with the average impacts previously shown in Table 1. In columns (2) and (3) of Table A9, we find that the labor-supply response of mothers at the extensive margin is much stronger in regions with higher coverage. In regions where daycare expanded more, the policy boosted maternal labor-force participation 67% more on average, well above the average effect of 7.8 percentage points in the entire province. In regions in the bottom tercile of the daycare coverage rate distribution, the increase in maternal employment is substantially lower and this estimate is statistically significant. High-coverage areas thus appear to be the regions that were driving most of the average impacts (reported in column 1), consistent with the positive correlation between expansion rates and policy impacts uncovered using the main specification.

In columns (4) to (6) of Table A9 and in Table A10, we examine the impact of the policy on other components of households' time allocation, namely hours worked by the mother and childcare use. For these outcomes as well, we find that using this indicator variable for low-expansion regions yields the same qualitative results as our more flexible specification of the treatment.

Impact on parenting practices. The policy increased maternal labor supply and thus mechanically reduced the time mothers spend at home with their children. If maternal care time and parenting effort are substitutes, we might expect mothers to compensate for the reduced time at home by spending more quality time with the child. In Table A11, we estimate the heterogeneous impact of the reform on the weekly frequency of reading to the child, our measure of parenting effort, by expansion status.

Consistent with previous results by Molnár (2023), we find that the policy had a positive impact on reading time at the bottom of the reading distribution. Point estimates suggest that parents were 4.4 percentage points more likely to read at least once per week and 5.4 percentage points less likely to never read to the child. We detect no short-run impact at the top of the reading distribution (reading daily). As for the time-allocation outcomes, the average impacts are driven by the most treated regions. For instance, the estimated decrease in the propensity to never read is almost entirely concentrated in high-expansion regions. These results thus suggest that mothers compensated for their increased work hours by exerting more effort parenting when they are home.

Robustness checks. We additionally perform a robustness check on these results. In Table A4, we compare our results to using only Ontario, the most similar province to Québec in terms of size and economy, as a control group. The estimates are quantitatively very similar to the previous specification.

A.5 Long-run impact on eligible children

This section assesses the long-run impact of the Québec reform on children's educational attainment and earnings in the early career. For individual i of age a born in province p observed in census year t, we estimate the following model:

$$\begin{aligned} Y_{iapt} = {} & \alpha_a + \alpha_p + \theta_1 C_t + \theta_2\, Q_i \times C_t \\ & + \sum_{a=21}^{36} \{\theta_{3,a}\, Age_a \times C_t + \theta_{4,a}\, Age_a \times Q_i + \beta_a\, Age_a \times Q_i \times C_t\} + X'_{iapt}\delta + \varepsilon_{iapt} \end{aligned} \tag{19}$$

where Y_{iapt} is educational attainment (completion of a given degree), $Q_i = 1$ is a dummy equal to 1 if the individual is born in Québec. C_t is an indicator of whether the individual is observed in the 2021 Census (= 0 if observed in the 2016 Census). α_a and α_p are age and province fixed effects, respectively. X_{iapt} is a vector of controls (gender, marital status, number of children). Parameters of interest is the vector β_a, which capture the intent-to-treat policy impact.

The results are reported in Figure A11. We find no evidence of negative effects on educational attainment of eligible children in the long-run. This pattern holds for each educational level, namely university, high school, and college/CEGEP completion. The results suggest a positive but statistically insignificant impact on completion of a university degree, the most comparable outcome across provinces, and no impact at lower levels. This null result is consistent with the long-run trends in educational attainment, which have been surprisingly parallel across Canadian provinces over decades (see Figure A10). In Appendix Figure A12, we also show that there is no discernible difference in educational attainment within Québec between high- and low-expansion regions.

Next, we investigate long-term effects on eligible children's earnings in their early career using tax files data (LAD 2010-2022). We use an event-study specification comparing earnings of children residing in Québec to children living in the rest of Canada born between 1988 and 1999. Specifically, for individual i residing in province p in birth cohort c, we estimate:

$$Y_{icp} = \alpha \; + \sum_{\substack{c=1988 \\ c\neq 1992}}^{1999} \{\beta_c\, D_c \times Q_p \; + \theta_c\, D_c\} \; + \omega\, S_i + \gamma_p \; + \eta_{icp} \tag{20}$$

where Y_{icp} is either labor income or net earnings averaged between age 22 and 25. S_i takes the value of one if the individual is female, D_c is a dummy equal to one for birth cohort c, and γ_p are province fixed effects. Parameters of interest are the β_c, which capture the dynamic policy effects and pre-trends. We plot these coefficients in Figure A13. For both outcomes, we find no differential trends among cohorts that were never eligible and insignificant positive effects on eligible cohorts. These findings suggest that we can safely reject large negative effects on children in the long-run despite the fast expansion of the childcare network.

In light of the body of evidence documenting the role of early-childhood circumstances for lifetime success (see Almond et al., 2018), we might expect the Québec policy to have long-run impacts on economic outcomes. However, several factors could explain the null effects we document. First, while Baker et al. (2019) find negative impacts on health and behavior, they obtain mixed evidence on cognitive skills. For example, they obtain insignificant impacts on some test score measures, but a positive impact on the mathematics component of PISA tests. Second, there is mixed evidence on the persistence of the short-term negative impacts on non-cognitive outcomes. While Baker et al.'s (2019) results suggest such persistence, Haeck et al. (2015, 2018, 2022) find that most negative impacts on children and parental behavior eventually fade away. Third, it is possible that compensating behavior of parents in their children's education (see Molnár, 2023) might have compensated for the impacts of daycare enrollment.

A.6 Measurement and predictions of variables

To estimate our structural model, we have to impute values for missing data on variables of interest. These include wages and non-labor income, childcare expenses, and quality of non-maternal care. We describe below how we measure these variables and impute values when missing.

Wages and non-labor income. To estimate the model, we need to assign wage offers and to measure non-labor income for every household. This includes predicting a wage offer for non-working mothers as well as imputing the wage rate and non-labor income when income information is missing. In the NLSCY, the person most knowledgeable about the child (PMK) reports wages (for both the PMK and the partner) as well as household income. Given the absence of policy impacts on fathers' labor supply, we treat the father's income as non-labor income from the mother's point of view. We thus measure non-labor income as the difference between the reported household income and the mother's labor earnings (wages and self-employment income). We estimate Mincer-type models to predict real wages and income for households with missing income data. Variables used for predictions are the age and number of siblings in the household, parents' age, education and immigration status, the size of the area of residence, and a set of Census Metropolitan Area (CMA) dummies to capture local labor market variation.

Childcare expenses. The NLSCY contains measures of childcare expenses in the last two cycles. Respondents report their weekly expenses on childcare in cycles 7 and 8. We follow Molnár (2023) and measure the hourly price of childcare by dividing weekly expenses by the number of hours in institutional care. We make sure to remove households who have a subsidized space, which is observed in those waves. We then obtain predictions for childcare expenses in pre-reform data using variables common to all waves. Variables used for predictions are the age and number of siblings in the household, parents' age, education and immigration status, and the size of the area of residence.

Non-maternal care quality. We measure non-maternal care quality by constructing an index from three survey questions available in cycles 3 and 4. These questions concern parents' satisfaction with the interactions the caregiver has with the child, how the caregiver praises the child, and the activities that stimulate learning. They are phrased as follows:

> *How often would you say your caregiver praises and encourages [CHILD'S NAME], and responds promptly when he/she needs help or comforting?*
> *(1) Never (2) Rarely (3) Sometimes (4) Often*
>
> *How often does your caregiver plan activities and use toys and other materials to help [CHILD'S NAME] learn new things?*
> *(1) Never (2) Rarely (3) Sometimes (4) Often*
>
> *How often does your caregiver encourage [CHILD'S NAME]'s language development by talking to him/her and asking questions, as well as using songs and stories for this purpose?*
> *(1) Never (2) Rarely (3) Sometimes (4) Often*

To build the index, we sum the indicators for the three questions. The index therefore takes values from 4 to 16. Those questions are not asked to families not using childcare so we perform similar imputations as for the other variables above. Again, variables used for predictions are the age and number of siblings in the household, parents' age, education and immigration status, and the size of the area of residence.

A.7 Youth crime

In this section, we investigate the robustness of our main MVPF results to including long-run costs of juvenile criminal activity. Baker et al. (2019) find that children exposed to the reform at a young age experienced long-lasting negative consequences on behavior and non-cognitive outcomes. In particular, their results suggest a positive impact of the policy on youth crime at ages 12-20.

To be sure, prevalent youth crimes are rather "benign" offences such as thefts of small amounts, mischiefs, breaking and entering, failures to appear in court, and cannabis possession. Through the lens of the MVPF framework, increased criminal behavior can impact welfare through two channels: additional costs to victims and productivity losses for offenders, which reduces the WTP for the policy, and additional costs on the police and criminal justice systems, which are negative fiscal externalities. To take into account these costs to society, we perform a back-of-the-envelope calculation using estimates of costs of crime reported in Cohen (2020). Since these costs appear many years after the enaction of the policy, we apply a discount factor of 3% following Hendren and Sprung-Keyser (2020). However, results are qualitatively robust if we do not discount.[44]

We focus on the estimates from the richest specification (Table 5, column 3) in Baker et al. (2019). They find an average increase in yearly youth criminal activity of 212 crimes per 100,000 inhabitants. Given that crime rates in Québec are very low, this figure represents a rise of 22% of the mean. The authors further break down the crimes into four categories: against persons (rise of 167 crimes), against property (rise of 342), drugs (rise of 99), and other convictions (rise of 239). Since cannabis possession is likely not very costly and now legal in Canada, we focus on the other three categories. For each category, we consider the costs of the most common crime, which are non-aggravated assaults, theft of less than $5,000, and failure to appear in court.

We first multiply the crime rates impacts by the population of exposed youth in each post-reform cohort considered in the original study to obtain the total yearly impact.[45] Second, we multiply this number by the victimization and offender productivity costs, which enter the WTP, and the government services costs (on the criminal justice system), which imply a fiscal externality. Third, for each post-reform cohort, we discount future costs to obtain the actualized value of increased youth crime. The results are reported in Table A12. We obtain that these impacts, both on the WTP and the fiscal externality, amount to about 20 million dollars. They are thus negligible compared to benefits stemming from mothers' earnings gains. As a result, the benchmark MVPF only slightly decreases to 1.37 (see Table A8).

A.8 Comparison of MVPF estimates with other criteria

There are recent debates in Economics on the use of the MVPF as a welfare criterion to evaluate social programs. In particular, García and Heckman (2022a,b) criticize the use of this metric arguing that the MVPF approach *(i)* abstracts from the deadweight loss of taxation and thus from the social marginal value of public expenditure, *(ii)* assumes a fixed government budget and is silent about policies that loosen the government's budget constraint, *(iii)* uses an arbitrary ratio, *(iv)* cannot rank all policies, and *(v)* interprets some welfare-improving policies as "money pumps". They suggest the use of an alternative criterion, namely the net social benefit (NSB). The NSB is simply the difference between the policy's benefits and the cost to society. García and Heckman (2022b) define the NSB as:

$$\text{NSB} = \text{Benefits} - \Omega(\text{Direct cost}) \tag{21}$$

where $\Omega()$ is a potentially non-linear function, which notably captures the deadweight loss of public expenditure. In practice, however, a linearity assumption on Ω is often made in the literature and we follow this approach in our comparative exercise below. We compare our MVPF estimates with the NSB and the standard cost-benefit ratio in Table A13.

Before discussing the results we note, however, that in a reply, Hendren and Sprung-Keyser (2022) show that those critics originate from a misconception about their welfare criterion and

[44] We obtain a WTP to avoid the committed crimes of $32.11 million and a fiscal externality of $30.83 million when we do not discount. Detailed results are available upon request.

[45] The population of Québec residents aged 12 to 20 years old was approximately 850,000 over the years considered for this analysis (Institut de la Statistique du Québec, 2022).

that, in several contexts, the MVPF may be preferable to the NSB. For example, one key advantage of the MVPF framework is that it does not assume how the government finances the policy while the standard deadweight loss of taxation assumes an arbitrary linear income tax rate. The MVPF framework evaluates welfare impacts of budget-neutral programs by comparing two MVPFs: the one of an expenditure policy to the one of a revenue-raising policy. On arguments *(iv)* and *(v)*, the MVPF approach identifies policies that pay for themselves and for which recipients have a positive willingness-to-pay as Pareto improvements (defined as an infinite MVPF, not as a negative one as point *(v)* states). It is thus true that one cannot rank among Pareto improvements, but the message here is that the government should implement all those policies (at no cost) so ranking them is obsolete. Last, a fair criticism of empirical welfare analysis in García and Heckman (2022b) is that, in reality, the welfare costs of raising public revenue (the function Ω in 21) are likely non-linear. Such non-linearities in the deadweight loss of public expenditure cannot be accounted for in the MVPF framework, but the critique also applies to other standard criteria for evaluating social programs considered by García and Heckman (2022b). Estimating non-linear welfare costs of raising public revenue is a promising avenue for future research, but is beyond the scope of this paper. We refer the reader to those papers for a more extensive discussion.

Our comparative exercise in Table A13 reveals that, for the benchmark estimator, the choice of criterion substantially affects social-welfare conclusions. When focusing on earnings gains only, we find that the cost-benefit ratio (CBR) is lower than one (0.97) and that the NSB is negative (-$87.1M), suggesting that the policy is not socially desirable. However, under the structural estimator, we find that all three criteria point at the same conclusion: benefits are larger than costs and the policy should be implemented under these criteria. Indeed, we obtain a cost-benefit ratio of 1.79 and a positive NSB of more than $2 billion under our structural estimator. Therefore, we find that using these alternative criteria reinforce our conclusion that omitting non-pecuniary gains for mothers and the non-marginal nature of the reform leads to a substantial underestimation of social-welfare gains. In fact, under the CBR and NSB criteria, the sufficient-statistic approach would lead one to conclude that the policy should not be adopted.

A.9 Appendix Figures

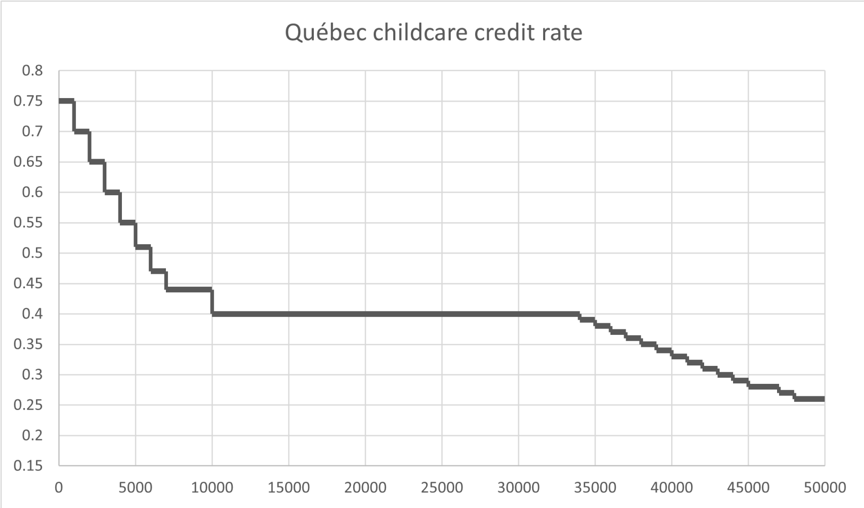


Figure A1: Refundable childcare credit rate by family income in Québec

Note: This Figure plots the refundable childcare credit rate as a function of family income. The refundable childcare credit was available to all families who used paid childcare before the 1997 daycare reform. After 1997, families using subsidized childcare were no longer eligible.

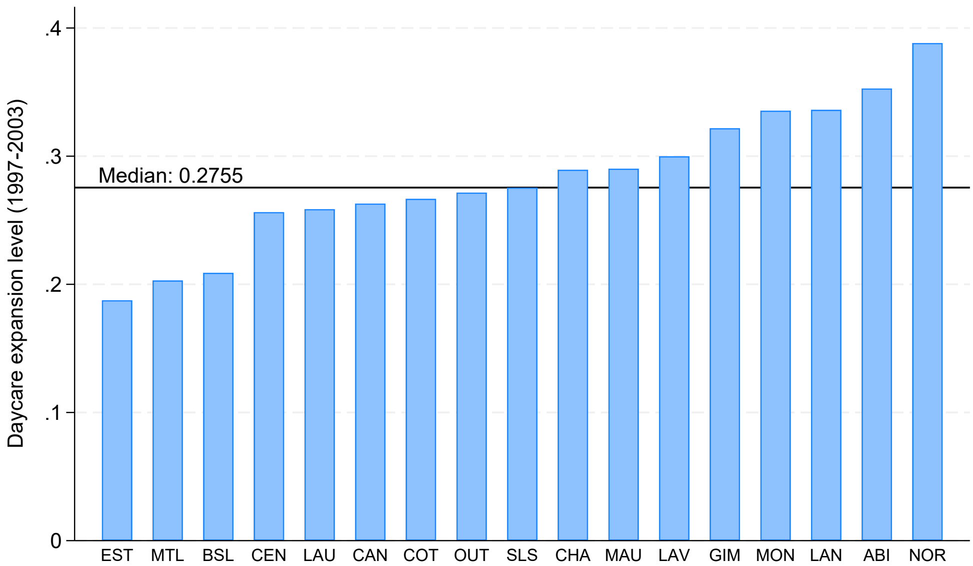
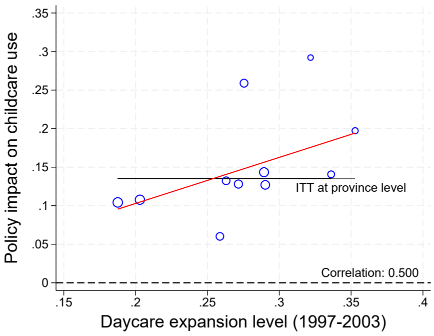
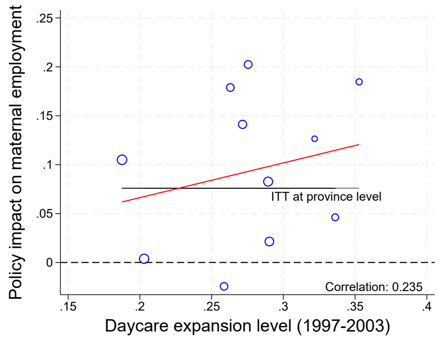


Figure A2: Distribution of the childcare expansion distribution across administrative regions

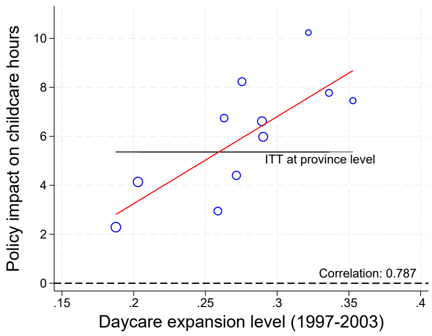
Note: This Figure plots the distribution of coverage expansions, measured as the difference in the childcare coverage rate between 1997 and 2003. It is constructed using data from a series of ministerial reports for daycare spaces (*Ministère de la Famille*, which held various other names), from the Québec Statistical Institute for the children population, and from authors' calculations. ABI = Abitibi-Témiscamingue; BSL = Bas-Saint-Laurent; CAN = Capitale-Nationale; CEN = Centre-du-Québec; CHA = Chaudière-Appalaches; COT = Côte-Nord; EST = Estrie; GIM = Gaspésie-Îles-de-la-Madeleine; LAN = Lanaudière; LAU = Laurentides; LAV = Laval; MAU = Mauricie; MON = Montérégie; MTL = Montréal; NOR = Nord-du-Québec; OUT = Outaouais; SLS = Saguenay-Lac-Saint-Jean



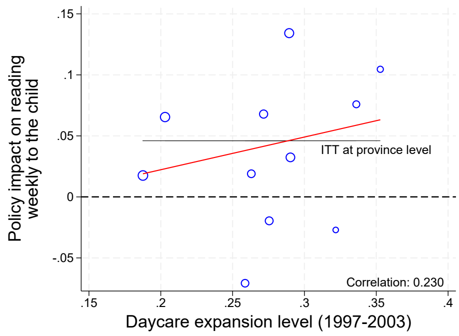
(a) Child in care



(b) Mother works



(c) Hours in care



(d) Reads weekly to child

Figure A3: Heterogeneous impacts of the Québec childcare reform and daycare expansion levels, controlling for rural areas

Note: These figures plot program effects on (a) childcare take-up (any non-parental care), (b) maternal employment, (c) childcare hours, and (d) reading weekly to the child estimated using equation (1) against the childcare expansion rate at the administrative-region level in regressions controlling for survey wave × rural area dummies. Circle sizes and the regression line are weighted by the number of observations. Included are regions with a sufficient sample size. The solid black line is the point estimate at the provincial level obtained from estimating our baseline specification, which we previously used to produce the ITT estimates in Table 1.

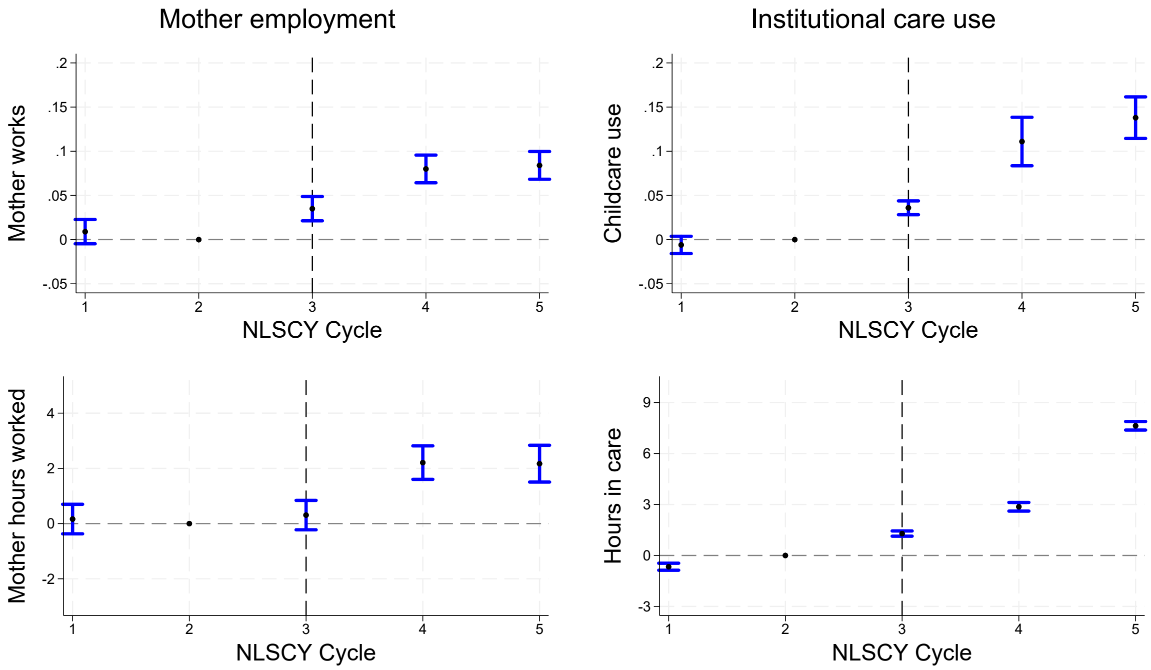
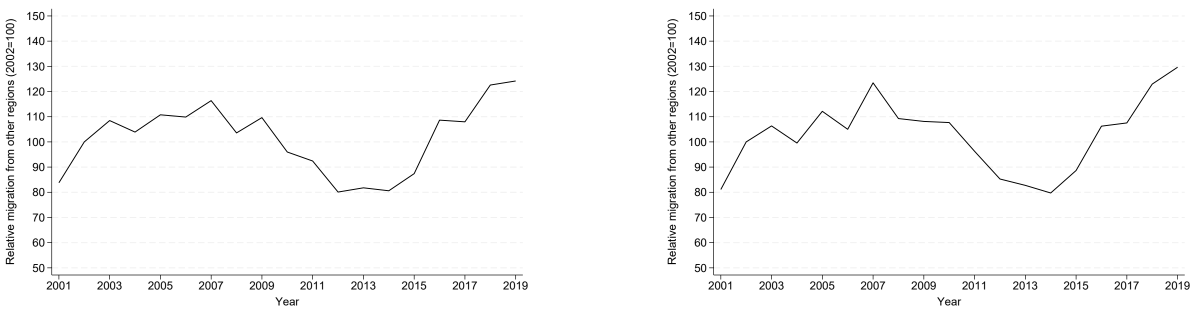


Figure A4: Dynamic impact of the Québec childcare reform on maternal labor supply and institutional care use

Note: These figures plot the coefficients of event-study regressions along with 95% confidence intervals. The data source is the first 5 waves of the NLSCY. Control variables are parents' age (in bins), age of the child, number and ages of siblings (in bins), population of the area of residence (in bins), education (both parents), and immigration status (both parents). The sample is restricted to two-parent families with a preschool-age child and with non-missing covariates.



Top tercile

Bottom tercile

Figure A5: Evolution of inter-regional migration by childcare expansion status

Data sources: Ministry of the Family for number of daycare spaces and Institut de la Statistique du Québec for population of preschoolers and net inter-regional migration.
Notes: These figures display the evolution of the net inter-regional migration flows of preschoolers (0-4 year olds) in two groups of regions relative to 2002 (normalized to 100). The left panel shows the changes in migration to regions which are part of the top third of regions who experienced the largest childcare expansion (measured as the increase in their coverage rate from 1997 to 2003). The right panel shows the equivalent time series for the bottom third.

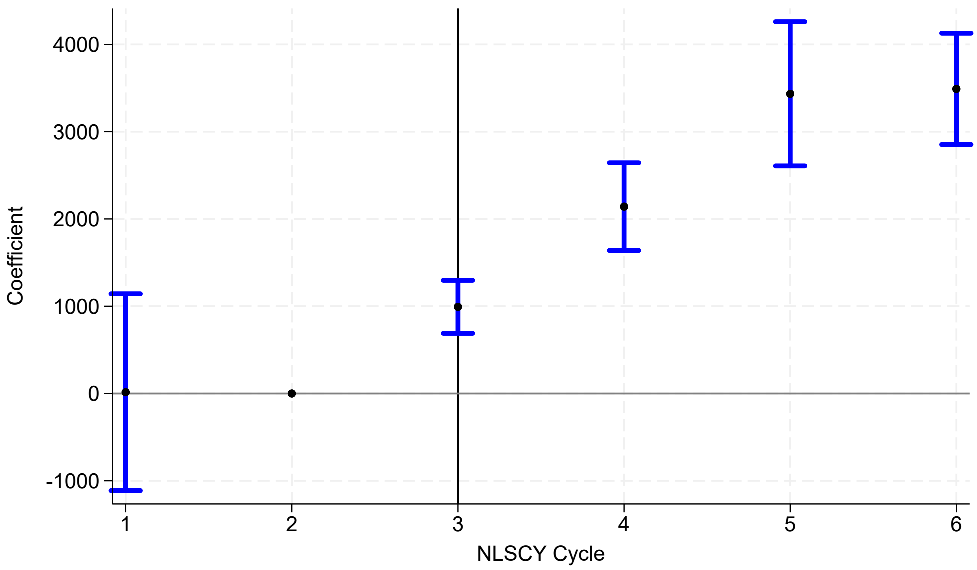
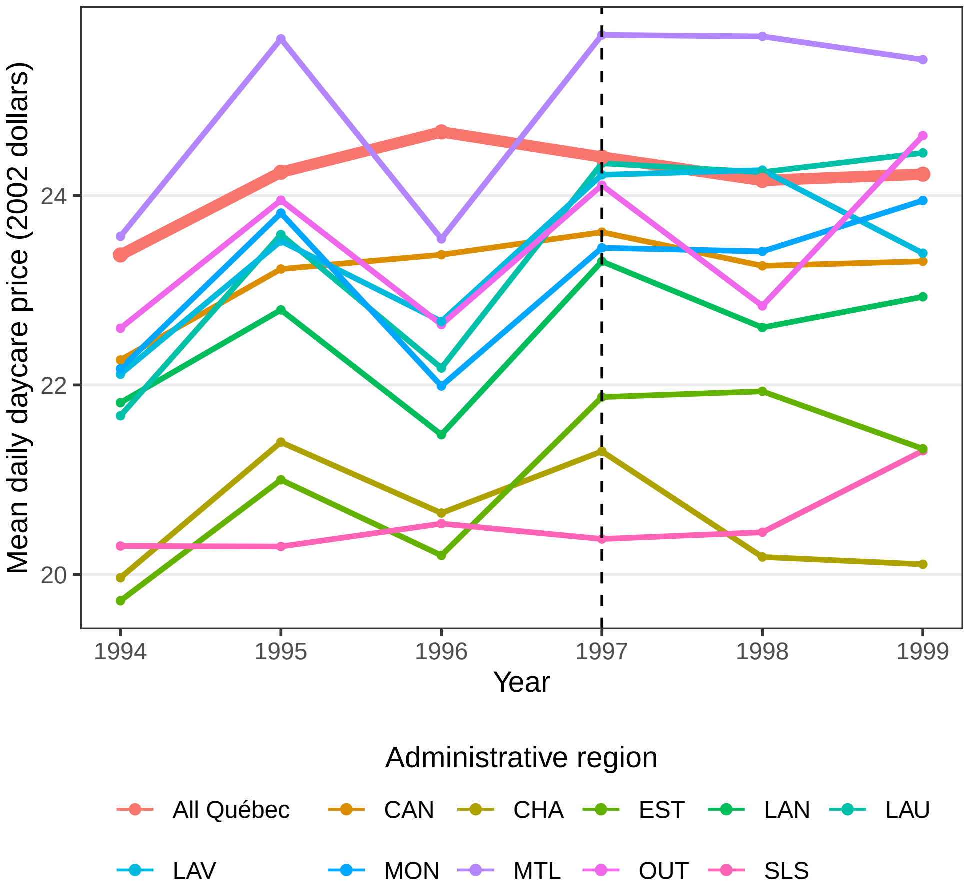


Figure A6: Dynamic impact of the Québec childcare reform on mothers' labor earnings

Note: These figures plot the coefficients of event-study regressions along with 95% confidence intervals. The data source is the first 5 waves of the NLSCY. Control variables are parents' age (in bins), age of the child, number and ages of siblings (in bins), population of the area of residence (in bins), education (both parents), and immigration status (both parents). The sample is restricted to two-parent families with a preschool-age child and with non-missing covariates.

Figure A7: Evolution of average daycare prices in unregulated network by administrative region



Data sources: Ministry of the Family
Notes: This figures plots the evolution of average daily daycare prices in constant 2002 dollars in selected administrative regions in Québec. The thickest line is the average in the entire Québec province. CAN = Capitale-Nationale; CHA = Chaudière-Appalaches; EST = Estrie; LAN = Lanaudière; LAU = Laurentides; LAV = Laval; MON = Montérégie; MTL = Montréal; OUT = Outaouais; SLS = Saguenay-Lac-Saint-Jean

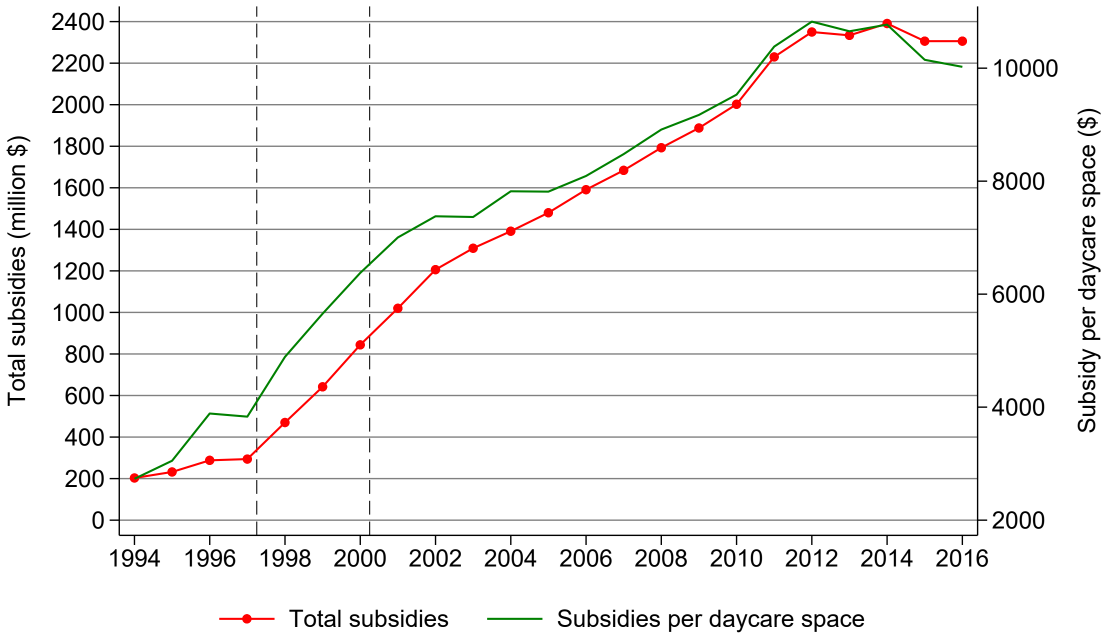
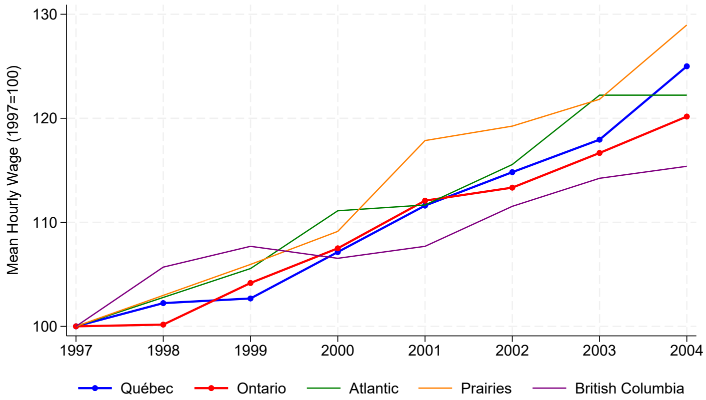
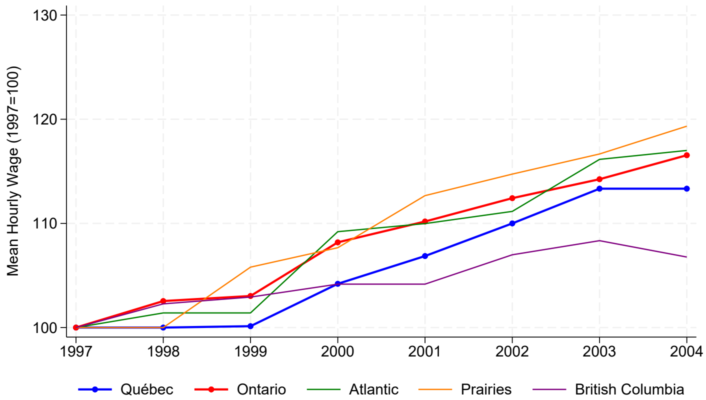


Figure A8: Evolution of daycare subsidies and subsidies per space in Québec

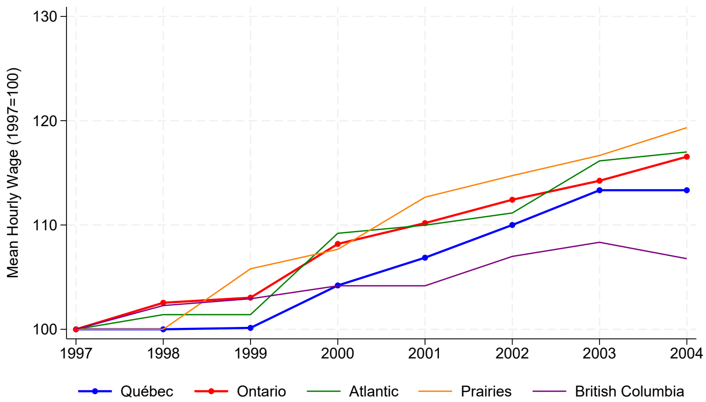
Note: This Figure displays the total subsidies to daycare facilities and families (red dotted line, left axis) and the subsidy per daycare space (green line, right axis). It is constructed using data from a series of budgetary reports of the Québec Treasury Board for the subsidy amounts and of the Ministry of the Family for daycare spaces. Additional costs of the program include additional administrative costs to operate the program, infrastructure subsidies to CPEs, and retirement pensions payments of daycare providers in CPEs. Those are nevertheless small in magnitude compared to direct subsidies.



(a) Childless women



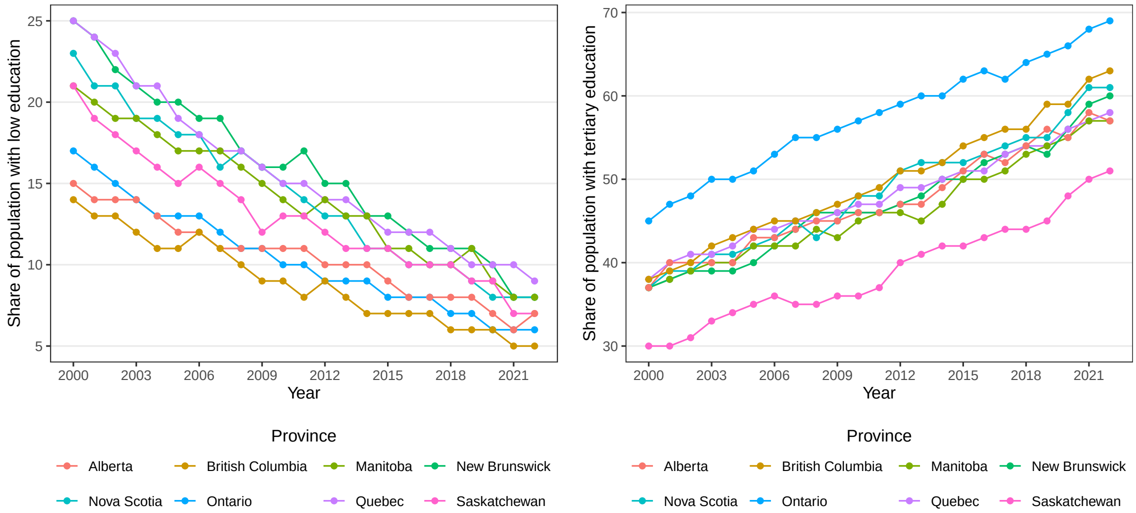
(b) Childless men



(c) All men

Figure A9: Trends in hourly wages across Canadian provinces, 1997-2004

Note: This figure plots average hourly wages of across Canadian provinces over time among childless women, childless men, and all men, respectively. The data source is the Canadian Labour Force Surveys 1997-2004. 1997 wages are normalized at 100.

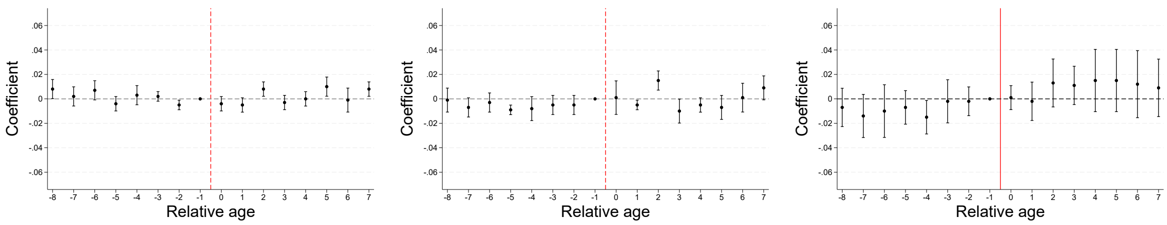


(a) No education

(b) Tertiary education

Figure A10: Long-term trends in educational attainment across Canadian provinces

Note: These figures plot population shares with (a) no high-school degree and (b) tertiary education in each Canadian province from 2000 to 2022. The data source is Statistics Canada Table 37-10-0130-01 from the Canadian Labour Force Surveys.



(a) High-school degree

(b) College degree

(c) University degree

Figure A11: Long-term effect of the Québec childcare reform on children's educational attainment

Note: These figures plot the regression coefficients on the triple interactions (β_a) from equation (19) using the 2016 and 2021 Canadian Census of population. The horizontal axis is the individual's relative age where the value of zero represents individuals aged 29 years old. Standard errors are clustered at the province level. 95% confidence intervals are reported in brackets.

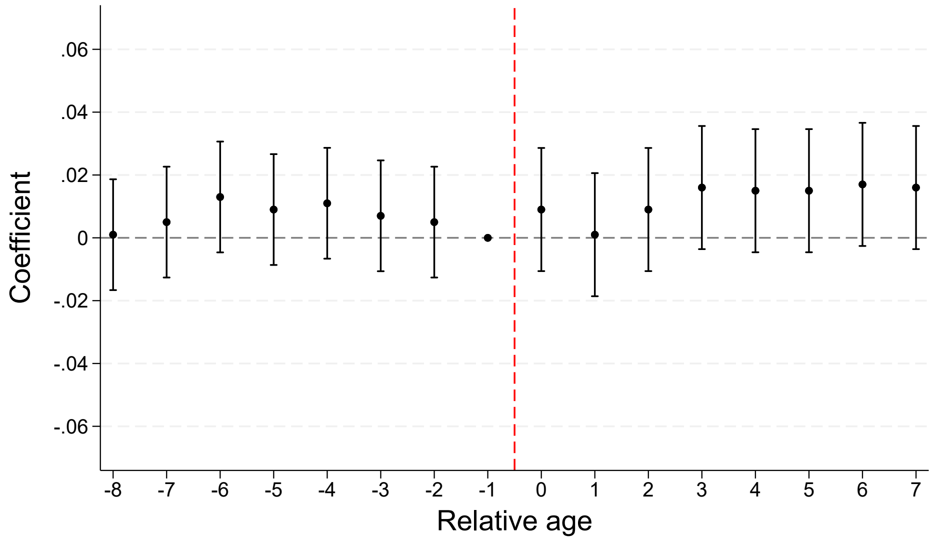
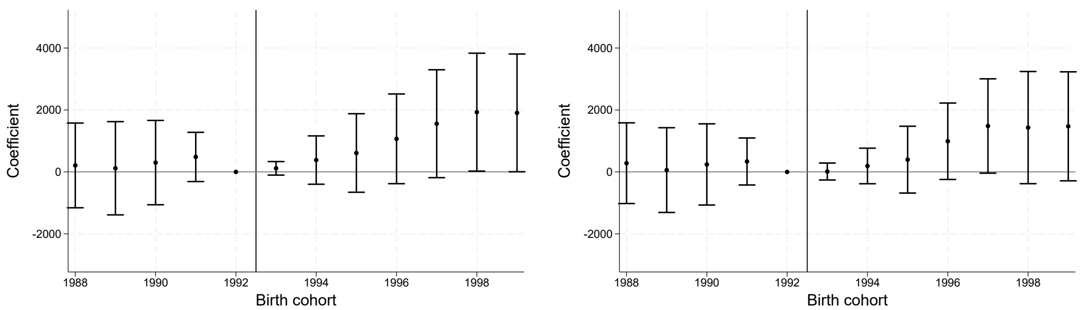


Figure A12: Long-run policy impact on children's educational attainment in low-expansion regions

Note: This figure plots the regression coefficients of event-study regressions using the 2016 and 2021 Canadian Census of population. The dependent variable is a dummy variable taking the value of one if the individual has completed high-school. The horizontal axis is the individual's relative age where the value of zero represents individuals aged 29 years old. 95 percent confidence intervals shown in brackets.



(a) Labor income

(b) Net earnings

Figure A13: Long-term effect of the Québec childcare reform on children's earnings

Note: These figures plot event-study coefficients from equation (20) using the Longitudinal Administrative Databank 2010-2022. The reduced-fee program ($5/day/child) began in September 1997 only for children aged 4 so children born in 1993 were eligible for only one year. All preschool-age children (0-4 years old) were eligible as of September 2000. Standard errors are clustered at the province level. 95% confidence intervals are reported in brackets.

A.10 Appendix Tables

Table A1: Descriptive statistics

	Québec		Rest of Canada	
	Pre	Post	Pre	Post
Panel A: household characteristics				
Age of mother	30.893	31.167	31.656	32.226
	(4.86)	(5.38)	(4.994)	(5.27)
Age of father	33.508	33.965	34.065	34.789
	(5.368)	(5.822)	(5.584)	(5.858)
Age of child	2.023	2.019	1.998	2.018
	(1.421)	(1.411)	(1.416)	(1.413)
Number of younger siblings	0.236	0.218	0.235	0.220
	(0.474)	(0.442)	(0.463)	(0.451)
Mother is immigrant	0.088	0.125	0.218	0.245
	(0.284)	(0.331)	(0.413)	(0.43)
Father is immigrant	0.096	0.126	0.209	0.239
	(0.295)	(0.332)	(0.406)	(0.427)
Number of older siblings	0.780	0.792	0.904	0.836
	(0.924)	(0.95)	(1.035)	(0.954)
Child is a girl	0.494	0.479	0.488	0.489
	(0.5)	(0.5)	(0.5)	(0.5)
Mother college degree	0.202	0.270	0.204	0.273
	(0.402)	(0.444)	(0.403)	(0.446)
Father college degree	0.195	0.239	0.215	0.262
	(0.396)	(0.427)	(0.411)	(0.44)
Mother no education	0.133	0.122	0.108	0.093
	(0.34)	(0.328)	(0.311)	(0.291)
Father no education	0.167	0.157	0.138	0.111
	(0.373)	(0.364)	(0.345)	(0.314)
Household lives in rural area	0.153	0.150	0.154	0.105
	(0.36)	(0.357)	(0.361)	(0.307)
Panel B: selected outcomes				
Child in care	0.418	0.630	0.407	0.482
	(0.493)	(0.483)	(0.491)	(0.5)
Hours in care	13.071	17.425	11.571	10.606
	(17.983)	(19.237)	(16.983)	(16.361)
Mother works	0.532	0.644	0.590	0.636
	(0.499)	(0.479)	(0.492)	(0.481)
Mother hours worked	17.541	21.575	17.929	20.250
	(18.176)	(17.982)	(17.84)	(18.448)
Father works	0.867	0.921	0.909	0.954
	(0.339)	(0.27)	(0.288)	(0.21)
Father hours worked	36.374	39.628	39.483	42.264
	(16.264)	(13.854)	(15.119)	(12.766)
Rarely/never reads	0.226	0.086	0.107	0.024
	(0.418)	(0.281)	(0.31)	(0.153)
Reads weekly	0.369	0.395	0.235	0.219
	(0.483)	(0.489)	(0.424)	(0.413)
Reads daily	0.379	0.462	0.645	0.740
	(0.485)	(0.499)	(0.479)	(0.439)
Child PPVT score	98.408	100.462	100.301	102.191
	(16.097)	(15.128)	(14.945)	(15.124)
Child MSD score	99.300	98.418	100.462	101.106
	(15.028)	(14.674)	(15.254)	(14.344)

Note: Pre-reform data are the first two waves (1994-1995 and 1996-1997) of the National Longitudinal Survey of Children and Youth (NLSCY). Post-reform data are waves 4 and 5 of the NLSCY (2000-2001 and 2002-2003). The sample contains 34,042 children aged 0-4 and their parents. Standard deviations are reported in parentheses.

Table A2: Impact of the Québec childcare reform on fathers' employment and institutional care use

Dep. var.:	Institutional care	Inst. care hours	Father works	Father's work hours
	(1)	(2)	(3)	(4)
Eligible_{pt}	0.131***	5.250***	0.005	0.159
	(0.009)	(0.240)	(0.010)	(0.523)
	[0.007]	[0.004]	[0.881]	[0.911]
N	33,575	33,320	34,012	31,497
R^2	0.069	0.076	0.161	0.09

Note: This table presents estimates from difference-in-differences regressions assessing the impact of the Québec childcare reform. Each columns refers to a different outcome. Specifically, we report estimates of the β coefficient from the following baseline specification estimated for mother i in province p in year t: $Y_{ipt} = \alpha + \beta \text{Elig}_{pt} + \gamma_p + \gamma_t + \delta X_{ipt} + \varepsilon_{ipt}$, where Elig_{pt} takes value 1 if the household resides in Québec after the reform. X_{ipt} is a vector of controls including maternal age, child age, number of siblings, population of the area of residence, parental education, parental immigration status, and provincial unemployment rates. γ_p and γ_t are province and survey year fixed effects. Standard errors clustered at the province level are displayed in parentheses. p-values computed using the Wild subcluster bootstrap of MacKinnon and Webb (2018) accounting for the small number of clusters are reported in brackets. The sample corresponds to the NLSCY sample following Baker et al. (2008), up to the 2003 survey wave. Level of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Heterogeneous impacts of the Québec childcare reform on mothers' employment, childcare use, and reading time

Administrative region	Child in care (1)	Hours in care (2)	Mother works (3)	Reading weekly (4)
Abitibi-Témiscamingue	0.205	7.589	0.193	0.103
	(0.065)	(2.521)	(0.028)	(0.028)
Bas-Saint-Laurent	0.191	6.551	0.341	0.232
	(0.041)	(1.185)	(0.078)	(0.053)
Capitale-Nationale	0.131	6.719	0.177	0.018
	(0.022)	(1.570)	(0.035)	(0.025)
Centre-du-Québec	-0.013	0.708	-0.187	0.021
	(0.039)	(0.993)	(0.064)	(0.025)
Chaudière-Appalaches	0.152	6.756	0.091	0.134
	(0.072)	(2.481)	(0.042)	(0.016)
Côte-Nord	0.331	10.279	0.285	0.115
	(0.010)	(0.816)	(0.015)	(0.015)
Estrie	0.111	2.474	0.111	0.016
	(0.084)	(3.654)	(0.021)	(0.035)
Gaspésie-Îles-de-la-M.	0.311	10.478	0.145	-0.029
	(0.051)	(1.500)	(0.045)	(0.042)
Lanaudière	0.142	7.765	0.048	0.076
	(0.017)	(0.590)	(0.033)	(0.016)
Laurentides	0.061	2.982	-0.025	-0.071
	(0.037)	(3.132)	(0.015)	(0.038)
Laval	0.192	5.745	0.014	0.074
	(0.020)	(0.920)	(0.044)	(0.065)
Mauricie	0.129	5.932	0.024	0.032
	(0.028)	(0.847)	(0.041)	(0.026)
Montérégie	0.137	5.660	0.106	0.025
	(0.039)	(1.609)	(0.052)	(0.041)
Montréal	0.108	4.176	0.004	0.065
	(0.046)	(2.297)	(0.036)	(0.051)
Outaouais	0.120	3.978	0.139	0.066
	(0.044)	(0.892)	(0.019)	(0.015)
Saguenay-Lac-St-Jean	0.261	8.173	0.206	-0.022
	(0.017)	(1.080)	(0.014)	(0.063)
R^2	0.119	0.113	0.109	0.056
N	33,709	30,915	33,758	33,171

Note: This Table reports program effects on childcare take-up (any non-parental care), institutional (formal) care use, maternal employment, and reading weekly to the child estimated using equation (1). The data source is waves 1-2-4-5 of the NLSCY. Control variables are parents' age (in bins), age of the child, number and ages of siblings (in bins), population of the area of residence (in bins), education (both parents), and immigration status (both parents). The sample is restricted to two-parent families with a preschool-age child and with non-missing covariates. Standard errors clustered at the province-year level in parentheses.

Table A4: Heterogeneous impacts of the Québec childcare reform on mothers' employment and childcare use by childcare expansion status, comparison with Ontario only

Dep. var.:	Mother works		Mother's work hours		Child in care		Childcare hours	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β_1 : Eligible$_{pt}$	0.076***	0.150***	1.850***	4.065***	0.147***	0.190***	5.481***	6.577***
	(0.002)	(0.015)	(0.138)	(0.656)	(0.033)	(0.035)	(1.547)	(0.946)
β_2 : Eligible$_{pt}$		-0.060***		-1.738**		-0.046**		-1.972
$\times$LowExp$_r$		(0.007)		(0.602)		(0.016)		(1.671)
Region (r) FE		✓		✓		✓		✓
r-level controls		✓		✓		✓		✓
p-value of $\beta_1 + \beta_2 = 0$		0.000		0.000		0.001		0.012
R^2	0.116	0.119	0.105	0.110	0.127	0.130	0.114	0.119
N	15,739	15,739	15,725	15,725	15,735	15,735	14,426	14,426

Note: This table presents estimates from difference-in-differences regressions assessing the impact of the Québec childcare reform using mothers residing in Ontario only as control. In odd columns, we report estimates of the β_1 coefficient from the following baseline specification estimated for mother i in province p in year t: $Y_{ipt} = \alpha + \beta_1 \text{Eligible}_{pt} + \gamma_p + \gamma_t + \delta X_{ipt} + \varepsilon_{ipt}$, where Eligible$_{pt}$ takes value 1 if the household resides in Québec after the reform. In even columns, the estimating equation is: $Y_{ipt} = \alpha + \beta_1 \text{Eligible}_{pt} + \beta_2 \text{Eligible}_{pt} \times \text{LowExp}_r + \gamma_p + \gamma_t + \gamma_r + \delta X_{ipt} + \varepsilon_{ipt}$ X_{ipt} is a vector of controls including maternal age, child age, number of siblings, population of the area of residence, parental education, parental immigration status, and provincial unemployment rates. γ_p and γ_t are province and survey year fixed effects. Standard errors clustered at the province level are displayed in parentheses. Level of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Pre-reform descriptive statistics by childcare expansion status

	Low exp.	High exp.
Panel A: Household characteristics		
Age of mother	30.625	31.104
	(4.55)	(5.062)
Age of father	33.889	33.261
	(5.815)	(5.022)
Age of child	1.981	2.060
	(1.42)	(1.424)
Number of younger siblings	0.242	0.233
	(0.484)	(0.467)
Number of older siblings	0.794	0.772
	(0.946)	(0.902)
Child is a girl	0.512	0.483
	(0.5)	(0.5)
Mother is immigrant	0.151	0.044
	(0.358)	(0.206)
Father is immigrant	0.162	0.050
	(0.369)	(0.219)
Mother college degree	0.192	0.212
	(0.394)	(0.409)
Father college degree	0.210	0.186
	(0.407)	(0.39)
Mother no education	0.174	0.104
	(0.38)	(0.305)
Father no education	0.192	0.149
	(0.394)	(0.356)
Household lives in rural area	0.131	0.165
	(0.337)	(0.371)
Panel B: Selected outcomes		
Child in care	0.417	0.422
	(0.493)	(0.494)
Hours in care	13.501	12.852
	(18.468)	(17.666)
Mother works	0.520	0.538
	(0.5)	(0.499)
Mother hours worked	17.122	17.798
	(18.355)	(18.053)
Father works	0.867	0.867
	(0.339)	(0.339)
Father hours worked	36.444	36.321
	(16.401)	(16.145)
Rarely/never reads	0.215	0.232
	(0.411)	(0.422)
Reads weekly	0.388	0.356
	(0.487)	(0.479)
Reads daily	0.373	0.384
	(0.484)	(0.487)
Child PPVT score	96.014	100.120
	(16.887)	(15.338)
Child MSD score	98.412	99.944
	(15.344)	(14.839)

Note: The data are the first two waves (1994-1995 and 1996-1997) of the National Longitudinal Survey of Children and Youth (NLSCY). Low-expansion regions are administrative regions (within Québec) in the bottom tercile of the childcare expansion distribution. The sample is restricted to two-parent families with a preschool-age child. Standard deviations are reported in parentheses.

Table A6: Determinants of local childcare expansions

	(1)	(2)
Initial coverage rate	-0.6756**	-0.2208
	(0.24)	(19.54)
Number of inhabitants (in 1,000s)		-0.0003
		(0.02)
Number of preschoolers (in 1,000s)		0.0055
		(0.39)
Share of medium-educated (high school)		0.0008
		(0.35)
Share of high-educated		0.0019
		(0.28)
Unemployment rate		1.0996
		(74.08)
Female unemployment rate		-1.0255
		(111.29)
Mean wages (in 1,000$)		0.0156
		(1.42)
Mean monthly rent ($)		-0.0005
		(0.03)
Constant	0.3903***	0.2479
	(0.04)	(2.12)
p-value of joint significance	0.0056	0.9999
R^2	0.352	0.515

Note: This table reports coefficients of linear regressions of childcare expansion rate (from 1997 to 2003) on the initial coverage rate and baseline regional characteristics. The data sources are a series of ministerial reports for daycare spaces (*Ministère de la Famille*, which held various other names) as well as the Québec Statistical Institute and the 1996 Canadian Census for the children population and the regional characteristics. The second-to-last row reports the p-value of the hypothesis that all the coefficients on baseline regional characteristics are jointly zero. Bootstrapped standard errors (1,000 replications) in parentheses. Level of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Child skill production technology parameters in different Canadian provinces

Parameter	Description	Québec	Ontario	Atlantic	Prairies	British Columbia
		(1)	(2)	(3)	(4)	(5)
δ_0	self-productivity	0.203***	0.177***	0.155***	0.191***	0.173***
		(0.017)	(0.014)	(0.015)	(0.014)	(0.024)
δ_e	maternal care	0.154***	0.130***	0.127***	0.061***	0.085*
		(0.026)	(0.025)	(0.027)	(0.026)	(0.048)
δ_d	non-maternal care	0.210***	0.083***	0.014	0.047	0.210***
		(0.033)	(0.029)	(0.032)	(0.062)	(0.066)
p-value of $\delta_e^{QC} - \delta_e^{p} = 0$			0.765	0.624	0.122	0.494
N		3,860	5,994	4,879	5,297	1,877
R^2		0.215	0.200	0.113	0.160	0.147

Note: This Table reports estimation results for the child human capital production function (equation 9) in different Canadian regions. The sample for Québec includes all pre-reform preschoolers with non-missing covariates. Bootstrapped standard errors in parentheses. Level of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Benchmark welfare estimate including costs of juvenile crime

MVPF components	Mean values	External sources used
Direct cost	\$2,617M	Québec Treasury Board
Willingness-to-pay		
Mothers of preschoolers	\$2,213M	CTaCS
Mothers of older children	\$1,102M	Lefebvre et al. (2009)
Taxes and reduced transfers	-\$856M	CTaCS
Youth crime (long-run)	-\$20.16M	Baker et al. (2019) and Cohen (2020)
Fiscal externality		
Tax returns and reduced transfers	\$856M	CTaCS
Youth crime (long-run)	-\$19.36M	Baker et al. (2019) and Cohen (2020)
MVPF	1.37	

Notes: This table outlines the components of the MVPF under the benchmark estimator including additional societal costs of increased youth criminal activity. The last column reports the external sources used for the policy's cost and other sources of fiscal externalities. The acronym CTaCS refers to the Canadian Tax and Credit Simulator of Milligan (2019).

Table A9: Heterogeneous impacts of the Québec childcare reform on mothers' employment by daycare expansion

Dep. var.:	Mother works			Mother's work hours		
	(1)	(2)	(3)	(4)	(5)	(6)
β_1 : Eligible$_{pt}$	0.078***	0.128***	0.156***	2.129***	3.667***	4.356***
	(0.007)	(0.018)	(0.014)	(0.298)	(1.000)	(0.654)
β_2 : Eligible$_{pt}$ $\times$LowExp$_r$		-0.053***	-0.063***		-1.770***	-1.751***
		(0.007)	(0.006)		(0.598)	(0.632)
LowExp$_r$		0.042***			1.189**	
		(0.006)			(0.588)	
Region (r) FE			✓			✓
r-level controls			✓			✓
Mean dep. var.	0.532			17.54		
p-value of $\beta_1 + \beta_2 = 0$		0.000	0.000		0.000	0.000
R^2	0.105	0.106	0.107	0.099	0.099	0.102
N	33,758	33,758	33,758	33,637	33,637	33,637

Note: Columns 1 and 4 report the average impacts previously shown in Table 1, while the remaining columns present estimates from regression equation (1). Control variables are parents' age (in bins), age of the child, number and ages of siblings (in bins), population of the area of residence (in bins), education (both parents), immigration status (both parents), and provincial unemployment rates. Region-level controls are shares of medium- and high-educated mothers and the number of preschoolers in the region r. The data source is waves 1-2-4-5 of the NLSCY. The sample is restricted to two-parent families with a preschool-age child and with non-missing covariates. Standard errors clustered at the province-year level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Heterogeneous impacts of the Québec childcare reform on childcare use by daycare expansion

Dep. var.:	Child in care			Childcare hours		
	(1)	(2)	(3)	(4)	(5)	(6)
β_1 : Eligible$_{pt}$	0.138***	0.164***	0.187***	5.736***	6.614***	7.325***
	(0.032)	(0.029)	(0.034)	(1.405)	(0.917)	(0.942)
β_2 : Eligible$_{pt}$ $\times$LowExp$_r$		-0.048***	-0.051***		-2.124	-2.276
		(0.014)	(0.016)		(1.443)	(1.556)
LowExp$_r$		0.018			1.182***	
		(0.013)			(0.151)	
Region (r) FE			✓			✓
r-level controls			✓			✓
Mean dep. var.	0.418			13.07		
p-value of $\beta_1 + \beta_2 = 0$		0.001	0.001		0.017	0.012
R^2	0.116	0.116	0.118	0.110	0.110	0.113
N	33,709	33,709	33,709	30,915	30,915	30,915

Note: Columns 1 and 4 report the average impacts previously shown in Table 1, while the remaining columns present estimates from regression equation (1). Control variables are parents' age (in bins), age of the child, number and ages of siblings (in bins), population of the area of residence (in bins), education (both parents), immigration status (both parents), and provincial unemployment rates. Region-level controls are shares of medium- and high-educated mothers and the number of preschoolers in the region r. The data source is waves 1-2-4-5 of the NLSCY. The sample is restricted to two-parent families with a preschool-age child and with non-missing covariates. Standard errors clustered at the province-year level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Heterogeneous impact of the Québec childcare reform on weekly frequency of reading to the child by daycare expansion

Dep. var.:	Rarely/never reads			Reads at least weekly			Reads daily		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_1 : Eligible$_{pt}$	-0.054***	-0.112***	-0.095***	0.044***	0.094***	0.046**	-0.015	-0.013	0.020
	(0.009)	(0.006)	(0.012)	(0.017)	(0.019)	(0.022)	(0.018)	(0.036)	(0.030)
β_2 : Eligible$_{pt}$ $\times$LowExp$_r$		0.083***	0.079***		-0.079***	-0.074***		0.006	-0.003
		(0.010)	(0.010)		(0.019)	(0.007)		(0.023)	(0.010)
LowExp$_r$		-0.051***			0.037***			0.027	
		(0.010)			(0.010)			(0.021)	
Region (r) FE			✓			✓			✓
r-level controls			✓			✓			✓
Mean dep. var.	0.226			0.748			0.379		
p of $\beta_1 + \beta_2 = 0$		0.084	0.107		0.678	0.410		0.752	0.520
N	33,171	33,171	33,171	33,171	33,171	33,171	33,171	33,171	33,171
R^2	0.170	0.170	0.171	0.053	0.053	0.056	0.165	0.165	0.168

Note: Columns 1, 4, and 7 report the average impacts previously shown in Table 1, while the remaining columns present estimates from regression equation (1). Control variables are parents' age (in bins), age of the child, number and ages of siblings (in bins), population of the area of residence (in bins), education (both parents), immigration status (both parents), and provincial unemployment rates. Region-level controls are shares of medium- and high-educated mothers and the number of preschoolers in the region r. The data source is waves 1-2-4-5 of the NLSCY. The sample is restricted to two-parent families with a preschool-age child and with non-missing covariates. Standard errors clustered at the province-year level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Costs of the Québec childcare reform from increased youth criminal activity

Type of crime	Victim costs	CJS costs	Offender productivity	Impact (BGM)	WTP	Fiscal externality
Persons (assaults)	$9,145	$3,594	$524	167 [59]	$17.71M	$6.58M
Property (theft)	$251	$1,922	$89	342 [93]	$1.28M	$7.21M
Drugs[1]	0	$4,523	$786	99 [29]	$0.85M	$4.91M
Other[2]	0	$176	$86	239 [54]	$0.32M	$0.65M
Total					**$20.16M**	**$19.36M**

Notes: Costs of crime estimates are taken from Cohen (2020) and are converted in 1997 Canadian dollars using the average exchange rate in 1997 (1.3252CAD/1USD). For each crime category, we use the crime most often committed by Canadian youth (in parentheses) as reported in Baker et al. (2019) (BGM). These cost estimates include crimes committed by adults, which are more costly on average, and should thus be interpreted as upper bounds. The WTP column should be interpreted as the WTP for avoiding the committed crimes and is the sum of the victim and offender productivity costs. The estimated policy impacts are taken from Table 5, column 3 of Baker et al. (2019). Standard errors are reported in brackets.

[1] The most prevalent drug crime is cannabis possession, but the data does not allow us to distinguish between drug possession and sale. These estimates are thus likely to be upper bounds.

[2] The most prevalent crime in the "other" category is failure to appear in court, but the data does not distinguish between types of "other non-traffic violations".

Table A13: Comparison between MVPF and alternative social welfare criteria

Criterion	Formula	Value
Benchmark estimator		
MVPF	Benefits / Net cost	1.40
NSB	Benefits - $(1+\phi)$ Cost	-$174.3M
CBR	Benefits / $[(1+\phi)$ Cost$]$	0.95
Structural estimator		
MVPF	Benefits / Net cost	3.56
NSB	Benefits - $(1+\phi)$ Cost	$2588.7M
CBR	Benefits / $[(1+\phi)$ Cost$]$	1.75

Note: We assume a deadweight loss of public expenditure $\phi = 1/3$ in this example, as in Hendren and Sprung-Keyser (2022). MVPF = marginal value of public funds, NSB = net social benefit, CBR = cost-benefit ratio

B Mathematical Appendix

B.1 Proof of equation (16)

First, let us show that we can write the indirect utility V as a function of policy parameters $(s(\theta), \pi(\theta))$ and exogenous variables only. Combining the time constraints (2) and (3) and substituting into the utility function, the Lagrangian of the agent's maximization problem is:

$$\begin{aligned}\mathscr{L} = &\gamma_c \ln(C) + \gamma_\ell \ln(T - T_c - L + T_d) + \gamma_h \ln(h_1) + \gamma_m \ln(T_c - T_d) - \gamma_{e,1}\, e\, (T_c - T_d)^{\gamma_{e,2}} \\ &- \gamma_{d,1} + \gamma_{d,2}\, \pi(\theta) \\ &+ \lambda \left[wL + I - C - (p - s(\theta))\, T_d\right] \\ &+ \mu \left[\ln(h_1) - H(T_d, e; h_0, q)\right]\end{aligned}$$

The first-order conditions for optimality at interior solutions are thus given by:

$$\frac{\gamma_c}{C} = \lambda \tag{22}$$

$$\frac{-\gamma_\ell}{T - T_c - L + T_d} = \lambda w \tag{23}$$

$$\frac{\gamma_\ell}{T - T_c - L + T_d} - \frac{\gamma_{T_m}}{T_c - T_d} + \gamma_{e,1}\gamma_{e,2} e (T - T_d)^{\gamma_{e,2}-1} = \lambda(p - s(\theta)) + \mu\, \frac{1}{T_c}\, (\delta_d \ln(q) - \delta_e \ln(e)) \tag{24}$$

$$-\gamma_{e,1}\, (T_c - T_d)^{\gamma_{e,2}} = \mu\, \frac{\delta_e (T_c - T_d)}{e\, T_c} \tag{25}$$

where λ is the marginal utility of income and μ is the Lagrange multiplier on the child development constraint. Rearranging terms in (23), we can write the optimal labor hours as a function of non-maternal care time T_d, and exogenous variables:

$$L^*(T_d, T, T_c, w) = \frac{\gamma_\ell}{\lambda w} + T - T_c + T_d \tag{26}$$

Substituting (26) into (24) yields:

$$-\lambda w + \frac{-\gamma_{T_m}}{T_c - T_d} + \gamma_{e,1}\gamma_{e,2} e (T - T_d)^{\gamma_{e,2}-1} = \lambda(p - s(\theta)) + \mu\, \frac{1}{T_c}\, (\delta_d \ln(q) - \delta_e \ln(e)) \tag{27}$$

Now, rearranging terms in (25), we can write the optimal parenting effort e as a function of non-maternal care time and exogenous variables:

$$e^*(T_d, T, T_c) = -\mu \frac{\delta_e}{\gamma_{e,1} T_c} (T_c - T_d)^{1-\gamma_{e,2}} \tag{28}$$

Therefore, substituting (28) into (27) and solving for T_d, we can express optimal non-maternal care as a function of exogenous variables and policy paremeter $s(\theta)$ only. Then, substituting this solution into (26) and (28), we get that optimal choices only depend on policy parameter $s(\theta)$ and exogenous variables $\omega = (T, T_c, w, p, q)$. Therefore, the indirect utility $V = \max_{C,L,T_d,e} U(C, L, T_d, e, \pi(\theta)) \quad s.t.\ (2), (3), (5), (14)$ solely depends on policy parameters $(s(\theta), \pi(\theta))$ and exogenous variables.

Now, we prove the result in equation (16). For an infinitesimal policy change, the difference in indirect utilities (the numerator in (15)) is simply the total derivative of $V(\theta_0)$ with respect to θ, which is given by:

$$\begin{aligned}\frac{dV(\theta_0)}{d\theta} &= \frac{V(\pi(\theta_0), s(\theta_0), \omega)}{d\theta} \\ &= \frac{\partial V}{\partial \pi}\frac{d\pi(\theta_0)}{d\theta} + \frac{\partial V}{\partial s}\frac{ds(\theta_0)}{d\theta}\end{aligned} \tag{29}$$

Applying the envelope theorem from the agent's maximization problem implies:

$$\begin{aligned}\frac{\partial V}{\partial \pi} &= \gamma_{d,2} \\ \frac{\partial V}{\partial s} &= \lambda\, T_d\end{aligned}$$

Substituting in (29), we obtain:

$$\frac{dV(\theta_0)}{d\theta} = \gamma_{d,2}\,\frac{d\pi(\theta_0)}{d\theta} + \lambda\, T_d \frac{ds(\theta_0)}{d\theta}$$

The second term in this equation is the dollar value of time in childcare. From the time constraints (2) and (3), we have that the time in childcare is a function of leisure and labor hours: $T_d = T_c - T + \ell + L$. Therefore, any change in childcare hours T_d induces a change in leisure or labor supply. In the budget constraint (14), however, the only other variable that relates to available time is labor hours. Thus, any increase in the left-hand-side of (14) due to an increase in available time must be due to an increase in labor hours. As a consequence, the dollar value of time in childcare must be equivalent to the dollar value of the mother's increased labor income. Replacing in the expression above thus yields the result:

$$\frac{dV(\theta_0)}{d\theta} = \gamma_{d,2}\,\frac{d\pi(\theta_0)}{d\theta} + \lambda\, w \frac{dL(\theta_0)}{d\theta}$$

B.2 Willingness to pay for a policy change

In this Appendix, we generalize the results highlighting the biases in the estimation of the willingness to pay for a large reform discussed in section 6.2. We use a Hendren (2016) framework slightly generalized so as to include non-pecuniary attributes. We first discuss the WTP of a single individual and then aggregation over all beneficiaries to move to social welfare.

B.2.1 Individual willingness-to-pay

Environment. Consider a decision maker $i \in \mathcal{I}$ facing the problem of choosing a vector of J market goods $\mathbf{x_i} = (x_i^1, ..., x_i^J)$, which can include labor-market activity, and K non-market choice variables $\mathbf{z_i} = (z_i^1, ..., z_i^K)$ to maximize a utility function $u_i(\mathbf{x_i}(\theta), \mathbf{z_i}(\theta))$ where θ is a set of policy choices of the government (the tax schedule, the level of public-good provision, the net transfers to the agent, and so on). The government's policy choice θ can potentially affect the agent's choices $\mathbf{x_i}$ and $\mathbf{z_i}$, her after-tax income y_i, and prices of goods. The agent faces a standard budget constraint as well as a set of additional constraints on the non-market choice variables. For simplicity, we suppose this set is a singleton so that the agent has to meet the budget constraint and one constraint on $\mathbf{z_i}$ (for example, a time allocation constraint). The decision problem thus writes:

$$\max_{\mathbf{x_i}(\theta),\, \mathbf{z_i}(\theta)} u_i(\mathbf{x_i}(\theta), \mathbf{z_i}(\theta)) \quad s.t. \;\; p(\theta) \cdot \mathbf{x_i}(\theta) \leq y_i(\theta)$$
$$g(\mathbf{z_i}(\theta)) = 0$$

where $p = (p_1, ..., p_J)$ is a price vector and g is differentiable in each of its arguments. Let $V_i(\theta) = U(\mathbf{x_i}^*(\theta), \mathbf{z_i}^*(\theta))$ be the agent's indirect utility under policy state θ.

The government now implements a policy change. The reform moves the policy state θ from the status-quo policy θ_0 to some new policy state θ_1. The agent's willingness-to-pay (WTP) for this policy change can be measured by the standard equivalent variation (EV), which we denote $\Delta y_i(\theta_0)$. That is, the WTP is the variation in income under θ_0 that would make the agent indifferent between the status quo and the new policy state:

$$\Delta y_i(\theta_0) = \frac{V_i(\theta_1) - V_i(\theta_0)}{\lambda_i} \tag{30}$$

where λ_i is the agent's marginal utility of income.

WTP for small policy changes. Let us consider first, as is the case with the sufficient-statistic approach, that the policy change is infinitesimal. For an infinitesimal (marginal) policy change (in θ), the numerator in (30), the difference in indirect utilities, is the total derivative of $V_i(\theta_0)$ with respect to θ. Under the additional standard assumption that prices of goods remain unchanged at the margin (i.e. assuming competitive markets for $\mathbf{x}$), we get:

$$\frac{dV_i(\theta_0)}{d\theta} = \lambda_i \, \frac{dy_i^*(\theta_0)}{d\theta} + \mu_i \, g'(\mathbf{z_i}^*(\theta_0)) \cdot \frac{d\mathbf{z_i}^*(\theta_0)}{d\theta} \tag{31}$$

where μ_i is the Lagrange multiplier on the second constraint.
Proof. The Lagrangian of the problem writes:

$$\mathscr{L} = U(\mathbf{x_i}(\theta), \mathbf{z_i}(\theta)) - \lambda_i \left[p(\theta)\mathbf{x_i}(\theta) - y_i(\theta)\right] - \mu_i g(\mathbf{z_i}(\theta)) \tag{32}$$

and thus the solution satisfies the first-order conditions:

$$\text{FOCs:} \quad \frac{\partial U_i(\mathbf{x_i}^*(\theta), \mathbf{z_i}(\theta))}{\partial \mathbf{x_i}} = \lambda_i \, p(\theta) \qquad \frac{\partial U_i(\mathbf{x_i}(\theta), \mathbf{z_i}^*(\theta))}{\partial \mathbf{z_i}} = \mu_i \, g'(\mathbf{z_i}^*(\theta)) \tag{33}$$

where λ_i is the agent's marginal utility of income and μ_i is the Lagrange multiplier on the second constraint. Let $V_i(\theta) = U_i(\mathbf{x_i}^*(\theta), \mathbf{z_i}^*(\theta))$ be the agent's indirect utility under policy state θ.

For an infinitesimal policy change, the difference in indirect utilities (the numerator in (30)) is simply the total derivative of $V_i(\theta_0)$ with respect to θ, which is given by:

$$\frac{dV_i(\theta_0)}{d\theta} = \frac{\partial U_i(\mathbf{x_i}^*(\theta_0), \mathbf{z_i}(\theta_0))}{\partial \mathbf{x_i}} \cdot \frac{\partial \mathbf{x_i}^*(\theta_0)}{\partial \theta} + \frac{\partial U_i(\mathbf{x_i}(\theta_0), \mathbf{z_i}^*(\theta_0))}{\partial \mathbf{z_i}} \cdot \frac{\partial \mathbf{z_i}^*(\theta_0)}{\partial \theta} \tag{34}$$

Using the first-order conditions (33), we have:

$$\frac{dV_i(\theta_0)}{d\theta} = \lambda_i \, p(\theta_0) \, \frac{\partial \mathbf{x_i}^*(\theta_0)}{\partial \theta} + \mu_i \, g'(\mathbf{z_i}^*(\theta)) \frac{\partial \mathbf{z_i}^*(\theta_0)}{\partial \theta} \tag{35}$$

Taking the derivative of the budget constraint with respect to θ yields:

$$\frac{\partial y_i(\theta)}{\partial \theta} = \frac{\partial p(\theta)}{\partial \theta} \, \mathbf{x_i}(\theta) + p(\theta) \, \frac{\partial \mathbf{x_i}(\theta)}{\partial \theta}$$

Therefore, assuming that prices are not impacted by the policy change at the margin (for example, assuming competitive markets for $\mathbf{x}$), the first term on the right-hand-side is null and we obtain that the impact of the policy on earnings is simply given by:

$$\frac{\partial y_i(\theta)}{\partial \theta} = p(\theta) \frac{\partial \mathbf{x_i}(\theta)}{\partial \theta} \tag{36}$$

that is, the additional spendings induced by the policy. Substituting (36) into (35) yields the result:

$$\frac{dV_i(\theta_0)}{d\theta} = \lambda_i \, \frac{dy_i^*(\theta_0)}{d\theta} + \mu_i \, g'(\mathbf{z_i}^*(\theta_0)) \cdot \frac{d\mathbf{z_i}^*(\theta_0)}{d\theta}$$

□

Therefore, the numerator of the WTP is the sum of two terms, the *pecuniary* benefits and the *non-pecuniary* gains stemming from the relaxation of the second constraint (e.g. the increase in available time). The intuition behind this result is the logic of the Envelope theorem, which implies that, at the margin, behavioral responses do not have a direct effect on utility. Thus, if one additionally assumes that non-pecuniary gains are negligible, which may be reasonable for marginal reforms, the difference in utilities boils down to the policy's impact on the beneficiaries' budget constraint. The WTP is then simply given by the causal effect of the policy on earnings ($\frac{dy_i(\theta)}{d\theta}$). This result is powerful because it implies that the treatment effect on beneficiaries' earnings is a sufficient statistic for the numerator of the MVPF (Hendren, 2016).

WTP for non-marginal reform. Consider now a discrete (or non-marginal) policy change. In this case, such as with the Québec childcare reform, the previous result does not hold anymore since Envelope conditions only apply to marginal reforms. In particular, behavioral responses, both for market and non-market choices, now have direct impacts on utility because the agent re-optimizes behavior. Moreover, for large reforms, non-pecuniary gains may be important. Thus, the treatment effect on earnings of beneficiaries is a biased estimate of the WTP.[46]

[46] To see this, suppose the utility function u is additively separable in $\mathbf{x_i}$ and $\mathbf{z_i}$ such that $u_i(\mathbf{x_i}, \mathbf{z_i}) = u_i^1(\mathbf{x_i}(\theta)) + u_i^2(\mathbf{z_i}(\theta))$. The difference in indirect utilities is given by:

$$V_i(\theta_1) - V_i(\theta_0) = u_i^1(\mathbf{x_i}^*(\theta_1)) - u_i^1(\mathbf{x_i}^*(\theta_0)) + u_i^2(\mathbf{z_i}^*(\theta_1)) - u_i^2(\mathbf{z_i}^*(\theta_0))$$

The first (resp. second) difference captures the overall impact of the policy on utility via behavioral changes in $\mathbf{x_i}$ (resp. $\mathbf{z_i}$). For non-marginal policies, differences in u_i^1 and u_i^2 are no longer the partial derivatives of $V_i(\theta_0)$. The *large-policy bias* is given by: $u_i^1(\mathbf{x_i}^*(\theta_1)) - u_i^1(\mathbf{x_i}^*(\theta_0)) - \lambda_i \, \frac{dy_i^*(\theta_0)}{d\theta}$. The bias stemming from the omission of non-pecuniary gains is simply $u_i^2(\mathbf{z_i}^*(\theta_1)) - u_i^2(\mathbf{z_i}^*(\theta_0))$ since these are ignored by assumption.

B.2.2 Social welfare

We now consider aggregation of individual beneficiaries' willingness-to-pay to obtain an estimate of the society's willingness-to-pay. Assuming there exists of a set of Pareto weights ψ_i, for each beneficiary i, social welfare at a given policy state θ is given by:

$$W(\theta) = \sum_{i \in \mathcal{I}} \psi_i V_i(\theta) \tag{37}$$

where W is the social welfare function and V_i is the indirect utility function of beneficiary i. This formulation is very general and can accommodate any social welfare function. It allows, for instance, social preferences for redistribution from richer to poorer individuals (Hendren, 2016).[47]

The society's WTP, which we denote by SWTP, for a reform is then given by (the monetary value of) the difference in social welfare between the the new (θ_1) and the status-quo (θ_0) policy states. Using the equivalent variation $\Delta y_i(\theta_0)$ as a measure of beneficiary i's WTP in dollars, the society's WTP is thus given by:

$$\text{SWTP} = \sum_{i \in \mathcal{I}} \psi_i \frac{V_i(\theta_1) - V_i(\theta_0)}{\lambda_i} \tag{38}$$

where we used equation (30). As Hendren and Sprung-Keyser (2020) note, the ratio ψ_i/λ_i is the marginal social utility of individual i's income.

Sufficient-statistic approach As argued in the previous section, if the reform and non-pecuniary gains are sufficiently small, the equivalent variation for a beneficiary boils down to the treatment effect on her earnings. As a naive estimator of the society's willingness-to-pay for a large reform, one can use this powerful result, as if the policy change were marginal. Therefore, using equation (31), this estimator can be written as:

$$\text{SWTP}_{marg} = \sum_{i \in \mathcal{I}} \psi_i \left(y_i^*(\theta_1) - y_i^*(\theta_0)\right).$$

Thus, if the social welfare criterion is prioritarian, this estimator is simply the sum of weighted pecuniary gains of all beneficiaries. In other words, the sufficient-statistic estimator of the SWTP is the treatment effect on (weighted) aggregate earnings.

Structural approach For discrete (large) policy changes, the Envelope theorem does not apply. Since beneficiaries do not simply react at the margin, behavioral responses have first-order impacts on utility. Therefore, one has to estimate the utility gains for each individual, which may include gains (losses) from re-optimization and non-pecuniary benefits. These can be estimated by simulating the policy in a well-specified structural model of behavior estimated on pre-policy data. To obtain the equivalent variation, one can use these estimates to find, for each beneficiary, the amount of additional income that would make her indifferent between the

[47] As additional examples, a utilitarian planner sets $\psi_i = 1 \; \forall i$ while a Rawlsian planner sets $\psi_{i'} > 0$ for i' such that $V_{i'} \leq V_i \; \forall i$ and $\psi_i = 0 \; \forall i \neq i'$.

extra cash and the implemented policy. This structural estimator of the SWTP is then the weighted sum of equivalent variations:

$$\text{SWTP}_{model} = \sum_{i \in \mathcal{I}} \psi_i \frac{\tilde{V}_i(\theta_1) - \tilde{V}_i(\theta_0)}{\tilde{\lambda}_i} \tag{39}$$

where $\tilde{w}$ denotes that variable w is estimated from the model.

C Survey of MVPF estimates

In this appendix, we assess the prevalence of using sufficient-statistic estimators to calculate social-welfare impacts of non-marginal policy changes in economics. To do so, we conduct our own survey of MVPF estimates appearing on the Policy Impacts Library of (Hendren et al., 2025). Of course, this exercise requires some judgment calls and we therefore only take the results of our survey as suggestive.[48]

We first have to define a criterion indicating whether a policy change can be considered as marginal. Importantly, envelope conditions allowing one to express the welfare effect of a policy change as a fiscal externality concern individuals' utility maximization problem. The "size" of the policy change should thus not be evaluated using the number of recipients (or scale) or as a function of how local is the treatment effect estimated, but rather by the size of behavioral responses at the *individual* level.

In line with our theoretical discussion above, we define a discrete (non-marginal) policy as one that induces significant behavioral re-optimization by recipients. For marginal changes, by the Envelope theorem, recipients do not re-optimize at the margin and only obtain utility gains from the small relaxation of constraints in their maximization problem. Ignoring non-pecuniary gains of small reforms may also be reasonable. However, large behavioral responses to a policy such as entry into the labor market suggest agents face a different economic environment and revise their optimal choices, suggesting the policy change is non-marginal. Omitting non-pecuniary gains (or losses) of abrupt changes in behavior may also lead to important biases in welfare estimates.

We have surveyed the first 24 papers appearing on the Policy Impacts Library webpage as of January 16, 2023. For each paper, in addition to providing basic information on the reform considered, we assess whether the policy change being studied satisfies the marginal-policy criterion defined above. We also list some potential non-pecuniary gains (or losses) omitted by the authors. Lastly, we briefly discuss whether the marginal-policy assumption is likely to affect the authors' welfare conclusions. In particular, we check whether papers finding apparently welfare-improving (resp. welfare-decreasing) policies are omitting utility losses (resp. gains) which biases their estimates upwards (resp. downwards).[49] In our discussions, we focus

[48] We note that some authors discuss the limitations of their analysis in similar terms. For most papers, the welfare analysis is not the main goal and authors may only see this exercise as illustrative of the economic returns of the reform studied.

[49] One of the stated advantage of the MVPF framework is that it "can be easily compared across programs" (Kline and Walters, 2016, p. 1815). However, for such comparisons across policy domains, which require a precise ranking of policies, obtaining a robust MVPF point estimate is crucial. Therefore, even if the biases do not affect the authors' general welfare conclusions (i.e. whether the policy is welfare-improving or not), omitting potentially

on policy impacts on outcomes studied by the authors (or closely related) and limit mentions to other fiscal externalities.

The detailed results are reported in Table C1. Two key findings stand out of this review. First, most papers who apply the MVPF framework do so in the context of a non-marginal policy change. Out of the 24 MVPF estimates studied, we find that at least 20 cases clearly do not satisfy our criteria for the policy studied being considered marginal. For example, in many cases, the policy is found to have large impacts on labor supply at the extensive margin. Such employment responses are not marginal, unless all individual entering the labor force were indifferent between working and staying out of the labor force in the status quo, which is unlikely.

Second, among those 20 cases, we argue that for at least seven, abstracting from the large nature of the policy change is likely to substantially affect the authors' welfare conclusions. For example, Giesecke and Jäger (2021) find an MVPF of 0.8 for the introduction of old-age pensions in the United Kingdom. This policy is likely not marginal given the impacts on labor supply and retirement decisions documented by the authors. For this reform, positive non-pecuniary gains for grandchildren such as reduced time taking care of the elderly as well as reduced financial stress related to retirement for beneficiaries are not included. If these gains are sufficiently large, the MVPF could be higher than one, suggesting that the policy is welfare-improving rather than being a net cost to society.

large utility gains (or losses) can be problematic for policy comparisons.

Table C1: Survey of MVPF estimates in Policy Impacts Library

Authors	**Causal estimates**	**Policy considered**	**Recipients**	**MVPF**	**Marginal?**	**Agents re-optimize?**	**Non-pecuniary gains/losses**	**Welfare conclusions likely affected?**
Duchini and Van Effenterre (2024)	Duchini and Van Effenterre (2024)	Introduction of primary school on Wednesday morning in France	Parents (mothers) of primary-school age children in France	3.6	No	Large adjustments in week schedule	Parenting time, more regular work schedules, child cognition	No. MVPF is already high and reduced child care contraints should generate even more welfare gains.
Deshpande and Mueller-Smith (2022)	Deshpande and Mueller-Smith (2022)	Conditioning of access to the Social Security Income on re-evaluation of disability at 18 years old	Disadvantaged youth who would have received guaranteed SSI in adulthood	-16.1	No	Impacts on crime found to be concentrated in income-generating criminal offenses. This may suggest normally sub-optimal decisions to catch-up for the income loss. Consistent with this, Deshpande (2016) finds positive impacts of SSI removal on labor supply.	Psychic costs (or benefits) of criminal activity	No. MVPF is high (in abs. value) and reduced income in youth likely to imply additional negative distortions.
Hendren and Sprung-Keyser (2020)	Zimmerman (2014)	Admission to Florida International University	Prospective students at FIU		No	Admission to university likely to induce significant changes in life trajectory, preferences, etc.	utility of attending school, effort costs, moving costs from other States	No. Even if benefits might be overestimated, the net costs are negative.

Authors	Causal estimates	Policy considered	Recipients	MVPF	Marginal?	Agents re-optimize?	Non-pecuniary gains/losses	Welfare conclusions likely affected?
Paradisi (2021)	Brenøe et al. (2020)	Maternal leave in Denmark (effect of having female employees close to giving birth)	Coworkers and mother's firm	0.74	Arguably yes	Loss of only one employee for a short amount of time. Firms found to compensate by marginal adjustments, but no decrease in labor demand.	Maybe change in working atmosphere	No. MVPF below 1 and no omitted positive impact.
Hendren and Sprung-Keyser (2020)	Currie and Gruber (1996); Cutler and Gruber (1996); Dave et al. (2015); Miller and Wherry (2019)	Medicaid Expansions to Pregnant Women and Infants in US States	Pregnant women and children in the USA	∞	No	Large employment responses of women and probably behavioral adjustments in health habits.	Health insurance likely an important source of stress for families.	No. The policy already pays for itself and there should not be significant losses missing.
Hendren and Sprung-Keyser (2020)	Cohodes and Goodman (2014)	Massachusetts' Adams (MA) Scholarship	High-scoring high-school students	0.72	No	Change college choice and less likely to graduate	-	No. MVPF below 1 and no omitted positive impact.
Kuka and Shenhav (2024)	Kuka and Shenhav (2024)	EITC Benefits to Recent Mothers (1993)	Low-income workers (new non-married mothers)	5.6	No	Large effects on employment and wages	Improved mother's health and their children's outcomes	No. MVPF is already high.

Authors	Causal estimates	Policy considered	Recipients	MVPF	Marginal?	Agents re-optimize?	Non-pecuniary gains/losses	Welfare conclusions likely affected?
Cabral and Dillender (2021)	Cabral and Dillender (2021)	Workers Compensation Benefit Generosity	Injured workers	0.46	No	Medical spending and benefit duration increase	Workers may value the increased medical spending and the consumption-smoothing benefits afforded by more generous coverage (acknowledged by the authors)	Yes. If these benefits are sufficiently high, MVPF could be higher.
Hendren and Sprung-Keyser (2020)	Cornwell et al. (2006)	Georgia HOPE Scholarship	High-school students graduating with at least a B average	4	No	Large effects on college enrollment	Utility of attending college, effort costs	No. MVPF is very high.
Bastian and Jones (2021)	Bastian and Jones (2021)	All post-1990 EITC Expansions for Women	Low-income workers (women)	3.18-4.23	No	Large effects on employment	Improved health, reduced crime, improved child outcomes. (All accounted for only as fiscal externalities.)	No. When accounting for all fiscal externalities, policy already "pays for itself".

Authors	Causal estimates	Policy considered	Recipients	MVPF	Marginal?	Agents re-optimize?	Non-pecuniary gains/losses	Welfare conclusions likely affected?
Giesecke and Jäger (2021)	Giesecke and Jäger (2021)	Introduction of Old-Age Pensions in the UK	Individuals age 70+	0.8	No	Re-optimization of work / retirement decision. Large monetary transfer (22% of average income)	Time to help taking care of grand-children. Large change in life schedule. Financial stress relief.	Yes. MVPF close to 1. Potential welfare gain for grand-children and relief stress.
Bergolo and Cruces (2021)	Bergolo and Cruces (2021)	Conditional Cash Transfers in Uruguay	Low-income households with children and/or pregnant women. Covering 42 percent of children under 18	0.61	No	Discrete change in behaviour (health and educational conditions). Change in take-up of other social programs	Parenting time, Costs or gains of complying to the program's health and educational requirements	Maybe. Depends if the "complying with conditionalities" aspect has a positive or negative value for individuals.
Wettstein (2020)	Wettstein (2020)	Introduction of Medicare Part D	Individuals age 65+	1.98	No	If full time workers are working to get the private insurance, the policy leads to a complete re-optimization	If agents are liquidity constrained, they may value drugs more than their price, which means the WTP is higher. Gain on health and released stress from not being able to buy drugs	Yes. Some positive effects are neglected. MVPF could be higher

Authors	Causal estimates	Policy considered	Recipients	MVPF	Marginal?	Agents re-optimize?	Non-pecuniary gains/losses	Welfare conclusions likely affected?
Ganimian et al. (2021)	Ganimian et al. (2021)	Early-Childhood Education in India	pre-schoolers	∞	Arguably yes	Not with the pilot. If the policy is enlarged to the whole population, it might change the population of recipients (Parents of lower or higher expected-gain children now enroll because they know that it will improve their kid's outcomes)	Schooling might be more pleasant.	No. MVPF already infinite, only additional gains.
Jácome (2022)	Jácome (2022)	Medicaid Eligibility for Teenagers in South Carolina	19 year olds losing the coverage	1.77 to 14.96	No	Induce significant life decision change. If some individuals have child to get access to the program, it change significantly their behaviours.	Reduced probability of being incarcerated. Reduced opportunities of earnings via crimes. Access to better jobs (better in non-paid compensations)	

Authors	Causal estimates	Policy considered	Recipients	MVPF	Marginal?	Agents re-optimize?	Non-pecuniary gains/losses	Welfare conclusions likely affected?
Deshpande et al. (2021)	Deshpande et al. (2021)	Different evaluation standards (at known age thresholds) for eligibility to the Social Security Income in the USA	3 groups defined by leniency of evaluation standards for SSI eligibility: ¿55, 50-55, ¡50 and spillovers to homeowners	1.04	No	Large impacts on tail events, events that occur infrequently and are associated with large drops in consumption (p.152) like bankruptcy and home sale	Most treated individuals are in great financial distress so potentially large relief and positive metal health impacts of becoming eligible to SSI.	Yes. MVPF around 1 and reduced financial distress and increased access to housing are likely significant gains
Bailey et al. (2024)	Bailey et al. (2024); Hoynes and Schanzenbach (2009, 2012); Hoynes et al. (2016)	Access to Food Stamps (average $4/person/day)	US families (parents and children)	56.25	Probably not	The policy reduced food insecurity, but these are necessity purchases that would happen anyway. However, the employment response of parents suggests that, without food insecurity, some parents would not be working.	Time parents can spend away from work (leisure or child care) thanks to the food vouchers	No. MVPF very high, only additional gains.
Kline and Walters (2016)	Kline and Walters (2016)	Head Start (targeted preschool program)	Disadvantaged children in the US	1.85-2.41	Yes, if focusing only on children, but parents likely to benefit as well.	Children do not choose to attend daycare so no omitted behavioral changes	The value of child development for parents, pareting time, etc.	Possible if negative impacts on parents' utility (eg: time spent with child reduced)

Authors	Causal estimates	Policy considered	Recipients	MVPF	Marginal?	Agents re-optimize?	Non-pecuniary gains/losses	Welfare conclusions likely affected?
Gray et al. (2023)	Gray et al. (2023)	Removal of work requirements for food stamps in Virginia	Potential beneficiaries of food stamps	0.86-1.15	No, which the authors acknowledge. They provide an MVPF estimate including a utility cost of working.	No strong employment responses, but removal of work requirements found to have large impacts on retention.	Contrary to Bailey et al. (2022), the authors consider utility costs of labor.	Yes. The authors themselves show that including utility costs of labor moves the MVPF from below to above 1.
Hyman (2018)	Hyman (2018)	Trade Adjustment Assistance program (retraining incentives and UI) for displaced workers	Displaced workers (due to shifts in production outside the USA)	1.14 in paper; 2.7 on Policy Impacts	No	Important employment responses to the reform. The author acknowledges the caveats of assuming Envelope conditions.	Utility costs of training programs	Possible. The MVPF estimate is slightly higher than one and some utility costs might be omitted.
Deshpande (2016)	Deshpande (2016)	Conditioning of access to the Social Security Income on re-evaluation of disability at 18 years old	Disadvantaged youth who would have received guaranteed SSI in adulthood	0.9 to 1.01	No	Changes in labor supply as a response to the income loss	Income stabilization value of SSI included in second estimate. However, utility cost of working (foregone leisure) is omitted.	Yes. When authors include the insurance value of stable SSI payments, MVPF goes from below to slightly above 1. The utility cost of foregone leisure might push the MVPF back below one.

Authors	**Causal estimates**	**Policy considered**	**Recipients**	**MVPF**	**Marginal?**	**Agents re-optimize?**	**Non-pecuniary gains/losses**	**Welfare conclusions likely affected?**
Finkelstein and Hendren (2020) and Baird et al. (2016)	Baird et al. (2016)	School-based deworming program (community-wide) in Kenya	Treated school-age children and spillovers to neighboring schools	∞	Arguably yes	Probably not. Children would probably go to school anyway if they were not sick.	-	No. MVPF is infinite and non-pecuniary losses likely small.
Cascio (2023)	Cascio (2023)	Targeted pre-K programs in the US	Preschoolers in US States	0 (no benefit)	No	Substantial evidence from the literature that pre-K programs yield positive economic returns in the long-run	The value of child development for parents, parenting time, etc.	Yes. There should be benefits for both parents and children.
Cascio (2023)	Cascio (2023)	Universal pre-K programs in the US	Preschoolers in US States	1.96-4.27	No	Substantial evidence from the literature that pre-K programs yield positive economic returns in the long-run	The value of child development for parents, parenting time, etc.	No. MVPF already high and mostly additional gains.

Appendix References

Bailey, M. J., Hoynes, H., Rossin-Slater, M., and Walker, R. (2024). Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence From the Food Stamps Program. *The Review of Economic Studies*, 91(3):1291–1330.

Baird, S., Hicks, J. H., Kremer, M., and Miguel, E. (2016). Worms at work: Long-run impacts of a child health investment. *The quarterly journal of economics*, 131(4):1637–1680.

Baker, M., Gruber, J., and Milligan, K. (2005). Universal childcare, maternal labor supply, and family well-being. Working Paper 11832, National Bureau of Economic Research.

Baker, M., Gruber, J., and Milligan, K. (2019). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy*, 11(3):1–26.

Bastian, J. E. and Jones, M. R. (2021). Do eitc expansions pay for themselves? effects on tax revenue and government transfers. *Journal of public economics*, 196:104355.

Bergolo, M. and Cruces, G. (2021). The anatomy of behavioral responses to social assistance when informal employment is high. *Journal of Public Economics*, 193:104313.

Brenøe, A. A., Canaan, S. P., Harmon, N. A., and Royer, H. N. (2020). Is parental leave costly for firms and coworkers? Working Paper 26622, National Bureau of Economic Research.

Cabral, M. and Dillender, M. (2021). The impact of benefit generosity on workers' compensation claims: Evidence and implications. Working paper.

Cascio, E. U. (2023). Does universal preschool hit the target? program access and preschool impacts. *Journal of Human Resources*, 58(1):1–42.

Cohen, M. A. (2020). *The costs of crime and justice*. Routledge.

Cohodes, S. R. and Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, 6(4):251–85.

Cornwell, C., Mustard, D. B., and Sridhar, D. J. (2006). The enrollment effects of merit-based financial aid: Evidence from georgia's hope program. *Journal of Labor Economics*, 24(4):761–786.

Currie, J. and Gruber, J. (1996). Health Insurance Eligibility, Utilization of Medical Care, and Child Health*. *The Quarterly Journal of Economics*, 111(2):431–466.

Cutler, D. M. and Gruber, J. (1996). Does Public Insurance Crowd out Private Insurance?*. *The Quarterly Journal of Economics*, 111(2):391–430.

Dave, D., Decker, S. L., Kaestner, R., and Simon, K. I. (2015). The effect of medicaid expansions in the late 1980s and early 1990s on the labor supply of pregnant women. *American Journal of Health Economics*, 1(2):165–193.

Deshpande, M. (2016). Does welfare inhibit success? the long-term effects of removing low-income youth from the disability rolls. *American Economic Review*, 106(11):3300–3330.

Deshpande, M., Gross, T., and Su, Y. (2021). Disability and distress: The effect of disability programs on financial outcomes. *American Economic Journal: Applied Economics*, 13(2):151–78.

Deshpande, M. and Mueller-Smith, M. (2022). Does welfare prevent crime? the criminal justice outcomes of youth removed from ssi. *The Quarterly Journal of Economics*, 137(4):2263–2307.

Drouin, C., Bigras, N., Fournier, C., Desrosiers, H., and Bernard, S. (2004). Grandir en qualité 2003. enquête québécoise sur la qualité des services de garde éducatifs. Technical report, Institut de la statistique du Québec.

Duchini, E. and Van Effenterre, C. (2024). School schedule and the gender pay gap. *Journal of Human Resources*, 59(4):1052–1089.

Finkelstein, A. and Hendren, N. (2020). Welfare analysis meets causal inference. *Journal of Economic Perspectives*, 34(4):146–67.

Ganimian, A. J., Muralidharan, K., and Walters, C. R. (2021). Augmenting state capacity for child development: Experimental evidence from india. Working Paper 28780, National Bureau of Economic Research.

García, J. L. and Heckman, J. J. (2022a). On criteria for evaluating social programs. Working Paper 30005, National Bureau of Economic Research.

García, J. L. and Heckman, J. J. (2022b). Three criteria for evaluating social programs. *Journal of Benefit-Cost Analysis*, 13(3):281–286.

Giesecke, M. and Jäger, P. (2021). Pension incentives and labor supply: Evidence from the introduction of universal old-age assistance in the uk. *Journal of Public Economics*, 203:104516.

Gray, C., Leive, A., Prager, E., Pukelis, K., and Zaki, M. (2023). Employed in a snap? the impact of work requirements on program participation and labor supply. *American Economic Journal: Economic Policy*, 15(1):306–41.

Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy*, 30(1):51–89.

Hendren, N. and Sprung-Keyser, B. (2020). A Unified Welfare Analysis of Government Policies*. *The Quarterly Journal of Economics*, 135(3):1209–1318.

Hendren, N. and Sprung-Keyser, B. (2022). The case for using the mvpf in empirical welfare analysis. Working Paper 30029, National Bureau of Economic Research.

Hendren, N., Sprung-Keyser, B., Stuart, E., et al. (2025). Policy impacts library. Last accessed: 2023-01-16.

Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long-run impacts of childhood access to the safety net. *American Economic Review*, 106(4):903–34.

Hoynes, H. W. and Schanzenbach, D. W. (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, 1(4):109–39.

Hoynes, H. W. and Schanzenbach, D. W. (2012). Work incentives and the food stamp program. *Journal of Public Economics*, 96(1):151–162.

Hyman, B. (2018). Can displaced labor be retrained? evidence from quasi-random assignment to trade adjustment assistance. Working Paper 3155386.

Institut de la Statistique du Québec (2022). Estimations de la population selon l'âge et le sexe, québec, 1er juillet 1971 à 2022. Last accessed 1 November 2023.

Jácome, E. (2022). Mental health and criminal involvement: Evidence from losing medicaid eligibility. Working papaer.

Japel, C., Tremblay, R. E., and Côté, S. (2005). Quality counts! assessing the quality of daycare services based on the quebec longitudinal study of child development. *Choices: Investing in Our Children*, 11(5).

Kline, P. and Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of head start*. *The Quarterly Journal of Economics*, 131(4):1795–1848.

Kuka, E. and Shenhav, N. (2024). Long-run effects of incentivizing work after childbirth. *American Economic Review*, 114(6):1692–1722.

Lefebvre, P., Merrigan, P., and Verstraete, M. (2009). Dynamic labour supply effects of childcare subsidies: Evidence from a canadian natural experiment on low-fee universal child care. *Labour Economics*, 16(5):490–502.

Miller, S. and Wherry, L. R. (2019). The long-term effects of early life medicaid coverage. *Journal of Human Resources*, 54(3):785–824.

Milligan, K. (2019). Canadian tax and credit simulator. Database, software and documentation, Version 2019-1.

Molnár, T. L. (2023). Costs of daycare, complementarities, and heterogeneous productivity of parenting time in child skill formation. Working paper.

Paradisi, M. (2021). Firms and policy incidence. Working paper.

Wettstein, G. (2020). Retirement lock and prescription drug insurance: Evidence from medicare part d. *American Economic Journal: Economic Policy*, 12(1):389–417.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711–754.