Optimal Intergenerational Sugar Tax: a Sufficient Statistics Approach

Kai Song

Optimal Intergenerational Sugar Tax: a Sufficient Statistics Approach

Kai Song

Abstract

Sugar over-consumption induces health problems, and the sugar-sweetened beverage (SSB) tax is debated worldwide. One of the motivations to tax SSB is "internality correction": correcting the over-consumption due to individual mistakes. Most existing models stop at a single generation, treating the government as the only agent exercising the correction. But in reality, parents are also correcting their children's behaviour: they provide food, form habits, and care about their children's welfare. Ignoring parental correction, the SSB tax may not be accurate. My thesis estimates the optimal SSB tax when considering the intergenerational impact the parents have on the children, extending the single-generation static tax to a multi-generation setting. I created a model where both the government and parents intervene; the optimal sugar tax is more than just adjusting parameters. I derive the sufficient statistics from the model to estimate the optimal SSB tax. I show empirically that accounting for intergenerational effect, the optimal SSB tax is between 8.8% to 12.2%, which is larger compared to the tax without intergenerational concern 5.4%.

Keywords: Public Economics, Optimal Taxation, Sugar-Sweetened Beverage (SSB), Health Economics, Internality

Acknowledgements

I want to express my gratitude to my supervisor, Dr Ash Craig, for his invaluable guidance and patience throughout this year. His insights and feedback have been instrumental in shaping the direction and quality of my work. I am grateful for his belief in me and my project, even when I was lacking in such.

I am also grateful to the faculty of the Research School of Economics at the Australian National University for their feedback. I have benefited greatly from those who are in the field of this Thesis: Associate Professor Maria Racionero, Associate Professor Yijuan Chen and Dr Elena Capatina, for their feedback that supported this research. I would also like to say thank you to Professor Simon Grant, Associate Professor Idione Meneghel, Associate Professor Evan Calford, Associate Professor Timothy Kam, Dr Timo Henckel, Dr Juergen Meinecke and Dr Motohiro Kumagai.

Finally, I would like to thank my friends for their support, tolerance, and care during this journey. Their support has been a constant source of strength.

Declaration

I declare that this thesis is my own original work and has not previously been submitted for assessment at this or any other university. All sources used have been appropriately acknowledged, and any assistance received in conducting the research or preparing the thesis has been disclosed.

Contents

1 Introduction

Over-consumption of added sugar is harmful to human health, contributing to diabetes, obesity and other health conditions.[1] A natural policy response to sugar over-consumption is to impose a tax on sugar or on specific products such as sugar-sweetened beverages (SSB). Up to August 2022, 105 national SSB taxes were in effect, covering approximately 51% of the world population (Hattersley and Mandeville, 2023). In Australia, the Parliamentary Budget Office (2024) proposed a 20% tax on the SSBs.

Economists justify the SSB tax in two ways: externality or "internality" correction. An externality arises when individuals do not bear the full cost of consuming SSB. For example, they may increase the burden on the public health system. Internality arises when individuals fail to correctly optimise their welfare due to misinformation, addiction or other behavioural bias (Allcott et al., 2019). Externalities are typically corrected by a "Pigouvian" tax: by setting the tax to its external marginal cost (Pigou, 1920). Correcting internality using "sin taxes", is more controversial, as it intervenes in personal choices, raising concerns about paternalism.[2] Internality provides an additional rationale for implementing SSB taxes, even in the absence of externalities.

I quantify the optimal SSB tax that incorporates both internally correcting and parental concern, which is important in SSB consumption.[3] First, I model a corrective tax that targets individual misoptimisation. Using the normative (V_t) and decision (U_t) utility, I distinguish between choices affected by mistakes and choices free from mistakes without overspecifying the behavioural model. Second, I derive the optimal corrective tax in terms of estimatable sufficient statistics. Finally, I use panel data from the National Longitudinal Surveys (NLS) to empirically estimate sufficient statistics. I identify those employed in health-related occupations as agents who are free from behavioural mistakes in sugar-sweetened beverage (SSB) consumption The estimated optimal tax decreases with the intergenerational habit impact, as

---

[1]See Della Corte et al. (2025); Gillespie et al. (2023); Janzi et al. (2024).

[2]Paternalistic government intervention has been widely debated in philosophy. See Dworkin (1972)'s work on the philosophical background of paternalism. Where he provided existing policy examples that are paternalistic in nature, and argues against the claim from Mill (1859): the government should only correct externality-generating activities. However, these debates offer limited traction for deriving the optimal sugar tax and lie beyond the scope of this thesis. Accordingly, I do not take a position on whether paternalism is normatively desirable. Instead, I adopt paternalism as a premise: given the possibility of a non-zero sugar tax, the analysis focuses on deriving the optimal internality-corrective tax under different assumptions.

[3]See Hu et al. (2020); Zhen et al. (2011).

stronger parental internalisation reduces the need for government correction. The empirical results showed that parents tend to be more biased compared to non-parents, suggesting an imposed direct parental impact that negatively affects the next generation's welfare. The simulated optimal SSB tax rate with intergenerational concern ranges from 8.8% to 12.0%, higher than the simulated optimal SSB tax rate without intergenerational concern 5.4%. If policymakers design sugar taxes without accounting for intergenerational impact, it would result in under-correction.

The recent literature addresses the SSB tax as a static instrument for correcting internalities and discusses its regressive property. The intergenerational impact of sugar is discussed separately without being incorporated into the SSB tax design. This thesis addresses this gap in the literature by extending the internality-corrective tax framework to incorporate the intergenerational impact of sugar.

Parents influence their children's behaviour, for example, through direct food provision, habit transmission, and within-household education. To capture this channel of influence, I model individual welfare as depending not only on one's own behaviour but also on the previous generation's consumption, so that parental decisions affect their children's welfare.

My model assumes that people make mistakes with regard to their sugar consumption. This may be because of a lack of information or a behavioural bias. I represent these mistakes by assuming that they maximise decision utility. This may differ from normative utility, which is what determines their level of well-being.[4] I extend the analysis to include parents as an additional source of correction. Parents' role introduces an intergenerational dimension to internality correction, as parental choices affect both current and future welfare outcomes.

I assume parents value both their own utility and their child's future welfare, scaled by an intergenerational discount factor that reflects how much they care about the next generation. However, parents may not perfectly perceive what truly benefits their child. If parents misoptimise their own welfare, they may also misrepresent their child's welfare. However, they may act with greater accuracy when making decisions on behalf of their children than the government would. Put differently, parents may be subject to behavioural bias when making choices for themselves but are relatively unbiased when deciding for their children.

To capture this variation in parental accuracy, I introduce a parameter $\alpha \in [0, 1]$, which measures how correctly parents perceive and act on their child's true welfare.

---

[4] I also considered traditional behavioural models, such as dynamic inconsistency, but excluded them due to data limitations and difficulties in parameter estimation.

Formally, individuals behave as if they are maximising

$$U_t + \delta(\alpha V_{t+1} + (1 - \alpha)U_{t+1}).$$

At one extreme, when $\alpha = 1$, parents fully understand how their consumption affects their child's future welfare. They still maximise their own decision utility but correctly internalise the intergenerational consequences of their choices. At the other extreme, when $\alpha = 0$, parents fail to recognise how their behaviour influences their child's welfare. Acting solely on their own biased preferences, they transmit their bias entirely to the next generation, amplifying welfare losses over time.

In reality, neither extreme is likely to hold. Rather than treating parental accuracy as discrete, I model it continuously with $\alpha \in [0, 1]$, where parents behave as if they are maximising an objective function that weighs these two perspectives according to their level of accuracy. A higher α indicates closer alignment with the child's true welfare, while a lower α reflects greater reliance on the parent's own biased decision framework.

The empirical objective is to identify different utility functions that describe behaviour with and without mistakes, with and without intergenerational concern. I group observations by whether they behave under mistakes and whether they are parents. Following Allcott et al. (2019), who used nutritionists and dietitians' behaviour as a normative benchmark, I identify individuals who are medical professionals as normative agents: those who behave as if they are maximising the normative utility. Their observed consumption serves as an empirical proxy for the counterfactual behaviour of fully informed individuals. Intergenerational concern is captured by an indicator for having at least one child in the household. Parents are assumed to maximise a joint utility function that values both their own welfare and their child's perceived future welfare, while non-parents maximise only their own.

I derive the price metric bias γ_1 and γ_0 to quantify the bias for parents and non-parents, which can be derived from the difference in log SSB consumption level between normative and general agents. The optimal tax formula can be written in terms of these two sufficient statistics. With intergenerational concern, an additional term proportional to the change in bias after becoming a parent $(\gamma_1 - \gamma_0)$ captures how parenthood may reduce or increase misoptimisation, and the optimal tax should account for such a change.

I use panel data from the National Longitudinal Survey of Youth 1979 (NLSY79) and the National Longitudinal Survey of Youth 1997 (NLSY97), both administered by the U.S. Bureau of Labor Statistics (2024a,b). This survey includes questions on

6

sugar-sweetened beverage consumption per week for multiple years.[5] The NLSY79 and NLSY97 complement each other in terms of the cohort's age coverage. The former tracks individuals born between 1957 and 1964, representing an older generation that entered adulthood in the late 1970s and 1980s, while the latter follows those born between 1980 and 1984, capturing a younger cohort transitioning into adulthood in the late 1990s and 2000s. One important but hard-to-measure aspect is the parameter η, which governs the impact of parental consumption of SSB on those for their child, therefore impacting the optimal SSB tax with intergenerational concern. The optimal tax rate ranges from 8.8% to 12% as η varies from 0 to 1. Future work is required to narrow down the range of what η is.

I extend the two-generation model of internality-corrective taxation to an infinite-horizon economy. This generalisation yields a sequence of stationary corrective taxes that internalises both present and inherited consumption biases. The stationary tax result converges to a fixed point, which is identical to the tax result in the two-generation model.

1.1   Related Literature

This thesis builds on a broad body of literature spanning multiple research areas. The following review outlines how these strands inform and motivate the framework developed here.

Sufficient statistics have become an important methodology in optimal tax analysis. It was first systematically reviewed by Chetty (2009) as a "bridge between the structural and reduced form methods". The sufficient-statistics approach begins from a simplified theoretical model whose parameters can be expressed in measurable quantities such as elasticities or covariances. Saez (2001) provides a canonical example, deriving optimal income tax rates using measurable elasticities of taxable income and income distribution parameters. I follow the sufficient statistic behavioural modelling methods proposed by Chetty (2009). Rather than assuming a specific behavioural model, Chetty starts from the notion of a "true" utility function that is free from bias and serves as the foundation for the social welfare function. Individuals, however, are not assumed to act as if maximising this true utility; their choices may deviate from it, creating a wedge between behaviour and welfare. This wedge justifies a corrective tax, whose magnitude can be empirically estimated from relationships

---

[5]NLSY79 included such a question in 2008, 2010, 2012, 2014, 2016, 2020, 2022 surveys. NLSY97 included such a question in 2009, 2010, 2011, and 2015 surveys.

7

between consumption and taxation.[6]

Economists have discussed internalities in two ways. The first model the decision mechanism and individual bias directly, and derive the correction for the mistake that offset the welfare loss. For example, O'Donoghue and Rabin (2006) sets up a "mistake" via dynamic inconsistency and derives the optimal tax. The second begins from a policy objective and derives welfare implications from it, given that the individual behaves with bias. For example, Chetty (2015) uses this method to discuss the policy of increasing savings. This thesis follows the first approach, modelling the mistake as a gap between decision and normative utilities, where the latter represents true welfare.

Normative and decision utility are widely used in behavioural public economics to model individual misoptimisation. Individuals behave as if maximising their decision utility, whereas an unbiased individual would behave as if they are maximising normative utility. The gap between the two provides the rationale for welfare-improving policy intervention. Bernheim and Rangel (2009) is one of the first works that established the gap between decision and normative utility, where he addressed the latter as true utility. The difference between normative and decision utility is often attributed to behavioural biases such as dynamic inconsistency.[7] These biases generate systematic gaps between actual and welfare-maximising behaviour, which is widely discussed in the behavioural public economics literature(O'Donoghue and Rabin, 2006; Gruber and Kőszegi, 2004; Lockwood, 2020).

There is some empirical work that estimates the magnitude of present bias in dynamic inconsistency, such as the work from Sadoff et al. (2019) on food choices.[8] This work uses an experiment to compare an individual's food choice when planning advance, and when shopping on the spot to recover the present bias.[9]

The gap between normative and decision utility can also arise from misinformation or limited knowledge. Instead of modelling knowledge, some empirical studies define the "normative behaviour" explicitly. Using observable variables—such as

---

[6]Following Chetty (2009), I assume quasi-linear utility in the biased commodity and a numeraire, with a fixed pre-tax price. These assumptions rule out income and general-equilibrium effects, and are discussed formally in the theoretical section.

[7]See Appendix B, where I show that dynamic inconsistency is a special case of the normative–decision utility framework.

[8]Also see Cheung et al. (2022)'s work on money, healthy and unhealthy food present bias among high-school students from high school students; and Augenblick et al. (2015)'s work on present bias for effort input when completing a task.

[9]This thesis does not adopt a present-bias framework because (i) experimental estimation is infeasible given data and time constraints, (ii) focusing on it would impose unnecessary structure and obscure other sources of internality, such as information or parental concern.

education or knowledge test results, one can estimate the gap through reduced-form analysis. Allcott and Taubinsky (2015) and Allcott et al. (2019) apply this method, using surveys and experiments to examine how information disclosure or knowledge affects consumer choices.[10] This abstraction increases theoretical simplicity and leaves room to derive the sufficient-statistic parameters.

The idea of taxing sugar dates back to Smith (1776), who described it as a commodity that is "almost universally consumed" but not one of the "necessaries of life", making it a suitable target for taxation (Cawley and Frisvold, 2023).[11]

An SSB tax can be justified only as a corrective instrument. If its purpose were solely to raise revenue, it would be undesirable, as it constitutes a differentiated tax. (Diamond and Mirrlees, 1971; Atkinson and Stiglitz, 1976) further demonstrate that differentiated commodity taxes are not optimal when preferences are homogeneous and weakly separable between labour and consumption, unless motivated by corrective (Pigou, 1920) or redistributive objectives. Yet sugar taxes are widely regarded as regressive (Sharma et al., 2014; Paraje et al., 2023); thus, redistribution cannot serve as a justification.

One might argue that SSB taxes are purely externality-corrective, yet existing designs often diverge from the true marginal externality. Most SSB taxes are volumetric rather than taxed by sugar content (Cawley and Frisvold, 2023). Some Middle Eastern countries even impose a 50% excise on all carbonated drinks, including diet soda, far above rates elsewhere.[12] Such discrepancies are difficult to justify under a purely externality-based rationale.

Recent advances in behavioural public economics reinterpret the SSB tax as an internality corrective tax in individual decision-making. O'Donoghue and Rabin (2006), Griffith et al. (2017), and Allcott et al. (2019) formalise this rationale theoretically and empirically. O'Donoghue and Rabin (2006) formally addresses the design of taxation to correct individual misoptimisation, where he used a dynamic inconsistency framework to set up an individual's (biased) choice, and then addressed the mistakes as the gap between dynamically consistent and inconsistent choices. He then assumed parameters, such as present bias, to simulate the optimal tax result. Griffith et al. (2017)'s work mainly focuses on the regressiveness of the SSB tax. In

---

[10]See also Attari et al. (2010), who compares public perceptions of energy savings with expert benchmarks.

[11]This early argument is limited, as welfare may arise from goods beyond mere necessities. Hence, Smith's reasoning reflects a paternalistic view rather than a neoclassical welfare argument.

[12]For comparison, the UK levies 18–24 pence per litre, while Philadelphia, USA, charges 1.5 cents per ounce.

her setting, she assumes the internality as an exogenous variable, and derives the optimal tax result with sufficient statistics on the redistribution component. My model is influenced by her and Chetty (2009)'s assumption regarding the utility function's formalities and the tax structure. Allcott et al. (2019)'s work is closely related to this thesis, where they quantified the internality using the normative and decision utility framework, as I have discussed prior. Furthermore, they used the nutritionist and dietitian's choice as an expert opinion, and assumed it to be free from bias and misinformation, and used elasticity and price to estimate such internality. I have adopted a similar approach: choosing the informed expert as the unbiased benchmark. I also extended their internality estimation from one generation to two generations.

While it remains debated whether sugar is addictive in a biomedical sense,[13] public health research has found important intergenerational patterns in food and SSB consumption. Rhodes et al. (2016) find strong parental effects on children's food consumption across diverse Australian households, while Talati et al. (2024) show that parents' SSB consumption is a strong predictor of their children's SSB intake. Economists emphasise its habit-forming nature and the lasting effects of childhood consumption. Hu et al. (2020) find that SSB regulations in California child-care centres reduced long-term sugar intake, and Zhen et al. (2011) show similar evidence of habit formation through declining long-run SSB tax revenue.

---

[13]See Westwater et al. (2016); Onaolapo et al. (2020); Westwater et al. (2018) for a medical perspective.

2   Theoretical Discussion

I constructed the Model of Intergenerational Sin Tax (MIST) and applied it in the context of sugar-sweetened beverages. The MIST is designed to capture three core mechanisms: the intergenerational impact of consumption behaviour, the mismatch between what individuals choose and what would be considered ideal, and the social planner's ability to influence outcomes through taxation.

2.1   Model Setup

Consider a discrete sequence of time. For any period $t \in \mathbb{N}$, there exist two types of agents: *children* and *adults*, while the children become adults in the period $t + 1$. In period t, each *adult* has one *child*. During childhood, the individual does not make any decisions and consumes a fixed share of the parents' consumption. For example, when parents consume the bundle (h, n) weekly, their children consumes $\xi(h, n)$ weekly, where $\xi \in (0, 1)$. For each adult, their decision is affected by their childhood consumption.

In a period t, the representative adult consumes two types of commodities, healthy food (h_t) and SSB (n_t), while their child consumes $\zeta \in (0, 1)$ of what their parent consumes[14]. The parameter ζ captures the idea that children largely inherit their parents' consumption patterns: what parents eat, children eat too, only in smaller quantities. It provides a simple way to model how dietary habits are transmitted across generations. The adult's consumption is subject to an exogenous income z, and the price of SSBs is p, assuming the price of h_t is 1.

A *taxation system* in period t, $(\tau_t, \mathcal{T}_t)$, is designed by a benevolent *social planner*, including two components: a linear commodity tax τ on n_t, and a lump-sum tax rebate $\mathcal{T}$. The commodity tax collection fully funds the tax rebate; hence, the social planner is subject to the fiscal feasibility constraint: $\mathcal{T}_t \leq \tau_t n_t$. Since $\mathcal{T}_t$ is a lump sum rebate, it is not distortionary for individuals, meaning the individual will not consume n_t to increase the rebate amount, and the rebate amount is not marginally sensitive to individuals' consumption.[15]

To capture the intergenerational impact, I included the consumption of the previous generation into the utility function. I further assume the utility function is

---

[14]The rest of the thesis considers the commodity h_t as a numeraire: a generalised commodity with its price equal to 1.

[15]This is a representative agent model. With the future extension regarding heterogeneity, however, it can be viewed as a non-distortionary redistributive transfer that preserves fiscal balance.

quasi-linear to simplify the analysis and remove the income effect. Hence, I define $U_t = h_t + u(n_t, n_{t-1})$, where $u(n_t, n_{t-1})$ is and monotonically increasing in n_t and twice differentiable in n_t, n_{t-1}. Similarly, I define the normative utility function as $V_t = h_t + v(n_t, n_{t-1})$, where $v(n_t, n_{t-1})$ monotonically increasing in n_t and twice differentiable in n_t, n_{t-1}. Since each child consumes a fixed share (ζ) of what their parents are consuming, incorporating parental consumption is equivalent to including one's own childhood consumption. The individual failed to maximise their normative utility due to behavioural bias, self-control issues, or misinformation.

2.2 Single-Generation Benchmark

For individuals in the single generation model $(t = 1)$, they behave as if they are maximising their decision utility, subject to a budget constraint affected by the taxation system. Formally:

$$\max_{h_1, n_1} U_1 = h_1 + u(n_1, n_0)$$
$$s.t. \ h_1 + (p + \tau_1)n_1 \leq z + \mathcal{T}_1 \tag{2.1}$$

Equation (2.1) is the individual budget constraint under the taxation system $(\tau, \mathcal{T})$. Suppose the optimiser for such an optimisation problem is $(h_1{}^*, n_1{}^*)$, the consumption at this level is called *satisfy individual rationality*. For social planners, they are maximising the normative utility, subject to individual rationality and fiscal feasibility. Formally:

$$\max_{\tau_1, \mathcal{T}_1} V_1 = h_1{}^* + v(n_1{}^*, n_0)$$
$$s.t. \ \tau_1 n_1^* \leq \mathcal{T}_1 \tag{2.2}$$

Equation (2.2) represents the fiscal feasibility constraint; when such a constraint is binding, the social planner allocates all tax revenue generated from SSB taxation to lump-sum rebates $(\mathcal{T}_t)$ within the same period.

Given that the utility function is monotonically increasing in n_1, the budget constraints in Equation (2.1) and (2.2) are binding[16]. Hence, the optimiser h_1^*, n_1^* satisfy both budget constraints described in Equation (2.1) and (2.2) at the binding case, therefore $h_1^* + pn_1^* = z$ needs to hold for the social planners.

To distinguish partial derivatives from time-indexed variables, this thesis uses the notation $f_{(1)}$ or $f_{(2)}$ to represent the first-order partial derivative, where the $f_{(1)}$

---

[16]Rest of the analysis in this thesis will write the budget constraint in the binding case, i.e. with "=" instead of "≤".

represents taking the partial derivative of the first argument in the function, and $f_{(2)}$ represents taking the partial derivative of the second argument in the function. Formally:

$$\frac{\partial f(a,b)}{\partial a} := f_{(1)} \qquad\qquad \frac{\partial f(a,b)}{\partial b} := f_{(2)}$$

Proposition 2.1. *For single period, the optimal commodity tax $\tau_1 = u_{(1)} - v_{(1)}$.*

Proving the Proposition 2.1 by finding the first order condition of τ_1 directly is complicated due to the unrevealed relationship between n_1^* and τ_1. Instead, I consider that the social planner is choosing the optimal h_1^*, n_1^*. I then recover the corresponding τ_1 using the solution of the consumer's problem.

Proof. Firstly, I address the social planner's problem of choosing h_1^* and n_1^*. At such an optimiser, the two budget constraints described in Equation (2.1) and (2.2) are both binding. Hence: $h_1^* + pn_1^* = z$. Hence, the Lagrangian is written as:

$$\mathcal{L} = h_1^* + v(n_1^*, n_0) - \lambda(h_1^* + pn_1^* - z) \tag{2.3}$$

The first order condition suggested:

$$\frac{\partial \mathcal{L}}{\partial h_1^*} = 1 - \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_1^*} = v_{(1)} - \lambda p = 0$$

Hence $\lambda = 1$ and $v_{(1)} = p$.

To identify what τ_1 induces such consumption, I address the consumer's problem by the corresponding Lagrangian:

$$\mathcal{L} = h_1 + u(n_1, n_0) - \lambda(h_1 + (p + \tau_1)n_1 - z - \mathcal{T}_1)$$

The first order condition suggested:

$$\frac{\partial \mathcal{L}}{\partial h_1} = 1 - \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_1} = u_{(1)} - \lambda(p + \tau_1) = 0$$

since $\mathcal{T}_1$ is lump-sum and independent of the individual n_1 choices: i.e. $d\mathcal{L}/d\mathcal{T}_1 = 0$.

Hence $\lambda = 1$, $u_{(1)} = p + \tau_1$. Since $p = v_{(1)}$, the τ_1 that induces h_1^* and n_1^* is $u_{(1)} - v_{(1)}$. $\qquad\square$

The proof of Proposition 2.1 illustrates the methodology applied throughout this thesis to derive the optimal SSB tax. The derivation proceeds in three steps. First, I reformulate the budget constraint of the social planner, from a budget constraint

13

about $(\tau, \mathcal{T})$, to one about (h_1^*, n_1^*) that jointly binds the consumer's and planner's conditions. Second, I derive the first order condition of the (h_1^*, n_1^*) for the social planner, suppose the social planner chooses (h_1^*, n_1^*) directly. Finally, I analyse the first order condition of the consumers, to recover the condition for τ, ensuring that (h_1^*, n_1^*) chosen by the social planner satisfies the individual rationality.

This expression of τ is referred to as *within-generation (WG) corrective*[17], since the result in Proposition 2.1 can be interpreted as "taxing the marginal internality", analogous to the externality-corrective tax from Pigou (1920). Formally, define the internality to be the gap between two types of utility, $I_1 = U_1 - V_1$, which equals $u(n_1, n_0) - v(n_1, n_0)$. The social planner in period one can tax the internality-generating commodity n_1, the optimal corrective tax in this setting equals the marginal internality with respect to n_1, equal to $\partial I_1 / \partial n_1 = u_{(1)} - v_{(1)}$.

The single-generation benchmark is the optimal tax when consumers are not concerned about the next generation, i.e. when there is no role for parents as a paternalistic agent. This is a relatively simple internality construction, and is how most of the existing literature addresses internality on sugar tax (Allcott et al., 2019; Griffith et al., 2017). The existing work that adopts this framework tends to focus on the regressiveness of the SSB tax and constructs a more complicated type space across consumers to capture the heterogeneity. While this thesis abstracts from the heterogeneity to obtain a clearer analytical expression, the following sections extend this benchmark to describe intergenerational linkages and parental paternalism as other factors that impact the optimal SSB tax.

2.3 Two-Generation Model

I extend the benchmark from a single period $(t = 1)$ to two periods $(t \in \{1, 2\})$. This introduces a second agent with the corrective motivation: the parent, whose decisions in $t = 1$ affect both their own welfare and the child's period 2 welfare. The social planner can choose 2 different tax systems as a *tax schedule*, namely $(\tau_t, \mathcal{T}_t)$, and $\mathcal{T}_t = n_t^* \tau_t$ for $t \in \{1, 2\}$. This ensures that, in each period, the social planner needs to rebate all the tax revenue generated by the SSB tax. Parents and social planners care about the next generation's welfare, and discount it by $\delta \in [0, 1]$. I assumed there is an intergenerational impact of the parental consumption of n_t on

---

[17]I address this as "within-generation (WG)" to differentiate from the "inter-generation (IG)". The "IG" term is soon to be mentioned in the two-generation model.

their children's consumption. Formally:

$$n_{t+1} = g(n_t), \ and \ g'(n_t) = \underbrace{\frac{dn_{t+1}}{dn_t}}_{IG \ habit}$$

g' captures the strength of the habit formation[18]. I address this term as *intergenerational (IG) habit.* The existence of such intergenerational influence is supported by studies in public health.[19]

2.3.1 The Optimal Tax

The second generation's consumer and social planners face the same problem as those in the single-generation benchmark: they cannot influence the previous generation's consumption and have no future generation to consider. Consequently, the optimal commodity tax in period 2 is:

$$\tau_2^* = u_{(1)} - v_{(1)}.$$

For the generation-one consumers, they are maximising an intergenerational objective function subject to their income, with $\alpha \in [0,1]$ representing how accurately parents perceive their child's true welfare, modelled as a weighting in maximising the child's normative utility. Consider the extreme cases when $\alpha = 0$ or 1. When $\alpha = 0$, the parents are uninformed about the next-generation's true welfare and behave as if they took the consideration in the discounted decision utility of the next generation. When $\alpha = 1$, the parents are informed about the next generation's true welfare and behave as if they are considering the discounted next generation's normative utility. Formally:

$$\max_{h_1, n_1} O_1 = h_1 + u(n_1, n_0) + \delta(h_2 + \alpha v(n_2, n_1) + (1 - \alpha)u(n_2, n_1))$$

$$s.t. \ h_1 + (p + \tau_1)n_1 = \frac{1}{1 + \zeta}(z + \mathcal{T}_1) \tag{2.4}$$

$$h_2 + (p + \tau_2^*)n_2 = z + \mathcal{T}_2^* \tag{2.5}$$

$$n_2 = g(n_1) \tag{2.6}$$

Equation (2.4) is the budget constraint faced by the generation-one consumers. The total spending on their own consumption is $1/(1 + \zeta)$ of the wealth from income and rebate, since their child are consuming the fixed share ζ of the parents'

---

[18]I acknowledge that this g' can be derived from the individual's utility function, and it will be demonstrated in the later sections. The notation g is used for simplicity and to increase the readability of the result.

[19]See Rhodes et al. (2016); Talati et al. (2024)

consumption. Hence, if the parents are spending $\frac{1}{1+\zeta}(z+\mathcal{T}_1)$ on their own consumption, then the total expenditure of parent and child is $z + \mathcal{T}_1$. Hence, Equation (2.4) is obtained. Equation (2.6) is the constraint on the habit formation: the parents have an impact on what the next generation is consuming and are aware of this when making their decisions regarding.[20]

Notate the optimiser as (h_1^*, n_1^*). The generation-one social planner's optimisation problem is to choose the τ_1, $\mathcal{T}_1$ to the decision (h_1^*, n_1^*), that maximises the social welfare function, subject to fiscal feasibility constraints. Formally:

$$\max_{\tau_1, \mathcal{T}_1} W_1 = h_1^* + v(n_1^*, n_0) + \delta(h_2^* + v(n_2^*, n*_1))$$

$$s.t. \ \tau_1 n_1^* = \frac{1}{1 + \zeta}\mathcal{T}_1 \tag{2.7}$$

$$n_2^* = g(n_1^*) \tag{2.8}$$

Equation (2.7) again assumes the government rebates all the tax revenue back to the consumers. Since the children are consuming the fixed share ζ of their parents' consumption, if the parents are consuming n_1 units of SSBs, the total consumption of parent and child on SSBs is $(1 + \zeta)n_1$, therefore generating the commodity tax revenue $(1 + \zeta)n_1\tau_1$. Equation (2.8) is identical to Equation (2.6) to capture the habit formation.

As discussed before, $\tau_2^* = u_{(1)} - v_{(1)}$, and $\mathcal{T}_2^* = n_2^*\tau_2^*$. I proposed the following proposition to describe the optimal commodity tax schedule (τ_1^*, τ_2^*).

Proposition 2.2. *The optimal tax commodity tax schedule (τ_1^*, τ_2^*) satisfy the following:*

$$\tau_1^* = \underbrace{u_{(1)} - v_{(1)}}_{WG \ correction} + \delta\{(1-\alpha)\underbrace{\overbrace{(u_{(2)} - v_{(2)})}^{IG \ direct \ correction} + \alpha\,[-\overbrace{\frac{dn_2}{dn_1}(u_{(1)} - v_{(1)})}^{IG \ habitual \ adjustment}]\}}_{IG \ correction}$$

$$\tau_2^* = \underbrace{u_{(1)} - v_{(1)}}_{WG \ correction}$$

where α is the level of knowledge on the next generation's normative utility. The optimal tax rebate schedule $(\mathcal{T}_1^*, \mathcal{T}_2^*)$ is $(n_1^*\tau_1^*, n_2^*\tau_2^*)$[21].

In the Proposition 2.2, the within-generation (WG) correction term shows up

---

[20]This is true since the parents are aware that the children maximise their decision utility, and I have implicitly assumed the parents and children have the same level of knowledge on the next generation's welfare, i.e. α is constant across generations.

[21]Refer to Appendix B for the detailed proof.

in the optimal tax formula for both generation one and generation two. For the generation-one optimal tax formulae, I address the addition term as *intergenerational (IG) correction*, which can be broken down into two new terms: I address the term

$$u_{(2)} - v_{(2)}$$

as the *intergenerational (IG) direct correction* term, and the term

$$- \underbrace{\frac{dn_2}{dn_1}}_{IG\ habit} \underbrace{(u_{(1)} - v_{(1)})}_{WG\ correction}$$

as the *intergenerational (IG) habitual adjustment* term: the negative product of the intergenerational habit formation effect and the within-generation correction. The further analysis focuses on τ_1^*, and I will discuss the economic interpretation of the new terms later.

The proof of Proposition 2.2 follows the same structure as Proposition 2.1 in the single-generation benchmark. I begin by combining Equations (2.4) and (2.7) so that the social planner's choice of taxation can be expressed in terms of optimal consumption. This yields the consolidated budget constraint:

$$h_1^* + pn_1^* = z$$

Since the second-generation planner faces an identical optimisation problem to the single-generation benchmark, the same condition holds for generation two.

$$h_2^* + pn_2^* = z$$

Next, I derive the first order condition of the social planner who chooses the optimal allocation (h_1^*, n_1^*), and the consumer, who chooses the consumption (h_1, n_1) given the τ_1. The τ_1^* is defined as the τ_1 at which the consumer's choice coincides with the social planner's optimal allocation.

2.3.2 Interpretation

To interpret the economic meaning of the terms *intergenerational direct correction* and *intergenerational habitual adjustment* terms, I proposed the following lemmas when α takes the value 0 or 1. This way, the terms are isolated and therefore clearer.

Lemma 2.1. *When $\alpha = 0$, the consumer maximise*

$$O_1 = h_1 + u(n_1, n_0) + \delta(h_2 + u(n_2, n_1)),$$

17

The optimal commodity tax schedule is

$$\tau_1^* = u_{(1)} - v_{(1)} + \delta(u_{(2)} - v_{(2)})$$

$$\tau_2^* = u_{(1)} - v_{(1)}$$

Lemma 2.2. *When $\alpha = 1$, the consumer maximise*

$$O_1 = h_1 + u(n_1, n_0) + \delta(h_2 + v(n_2, n_1)),$$

The optimal commodity tax schedule is

$$\tau_1^* = (u_{(1)} - v_{(1)}) - \delta \frac{dn_2}{dn_1}(u_{(1)} - v_{(1)})$$

$$\tau_2^* = u_{(1)} - v_{(1)}$$

One can make the following three observations from Proposition 2.2. First, the IG corrective component is the weighted average of the IG direct correction and the IG habitual adjustment terms. A lower α places more weight on the IG direct correction, and less on the IG habitual adjustment. A higher α (closer to 1) has the opposite effect. Second, the IG habitual adjustment term is negative. This implies that the more this term is weighted, the lower the optimal tax should be. Finally, the relative magnitudes of these two components, weighted by α, determine whether the intergenerational effect enlarges or diminishes the optimal tax relative to the single-generation benchmark.

The role of α in the Optimal Tax:

These observations are consistent with the definition of α as the degree to which consumers correctly perceive the next generation's normative utility. This consistency can be demonstrated more formally through the two extreme cases defined in Lemma 2.1 and Lemma 2.2. Consider Lemma 2.1 when $\alpha = 0$, the misalignment between the consumers and the social planner comes from both within-generation and inter-generation. Similar to the analysis of the single-generation benchmark result, the internality and marginal internality are as follows.

$$I_1 = O_1 - W_1 = u(n_1, n_0) - v(n_1, n_0) + \delta(u(n_2, n_1) - v(n_2, n_1))$$

$$\frac{\partial I_1}{\partial n_1} = u_{(1)} - v_{(1)} + \delta(u_{(2)} - v_{(2)})$$

The term $u_{(2)} - v_{(2)}$ corrects the mistakes parents make that directly affect their children through their own consumption n_1, and is not associated with the children's subsequent consumption of n_2. This includes mistakes such as sugar-induced illness and chronic health conditions, and does not include mistakes such as the next generation's over-consumption of sugar due to childhood habits.

Why the optimal tax may be lower?

To explain why the IG habitual adjustment term is negative, I analyse the Lemma 2.2 (where $\alpha = 1$) further. In contrast to Lemma 2.1, the misalignment between consumers and the social planner solely comes from within-generation. Hence, the internality and marginal internality are identical to the single-generation benchmark: with $I_1 = u(n_1, n_0) - v(n_1, n_0)$ and $\frac{\partial I_1}{\partial n_1} = u_{(1)} - v_{(1)}$.

The optimal commodity tax is not simply marginal internality, but rather the marginal internality, adjusted by the IG habitual adjustment term. This occurs since both the first-generation parents and the second-generation social planner act as corrective agents toward the second-generation consumers. Setting τ_1 equal to the marginal internality would over-correct the second generation's behaviour, since both the parents and the social planner are already correcting it. To prevent over-correction, the generation-one social planner reduces the tax rate.

How does parental behaviour offsets internality?

Consider the consumer's problem in Lemma 2.1 and Lemma 2.2. Intuitively, IG habitual adjustment captures the internality correcting passing through habit. Indeed, since $u_{(1)} - v_{(1)}$ measures the generation-two consumers' marginal internality. When scaled by IG habits, it captures by how much such correction has happened through habit.

For the first-generation consumers with $\alpha = 0$, the first order condition suggested:

$$u_{(1)} + \delta u_{(2)} - \delta \frac{dn_2}{dn_1}(u_{(1)} + (p + \tau_2^*)) = p + \tau_1^{\alpha=0}$$

For the first-generation consumers with $\alpha = 1$, the first order suggested:

$$u_{(1)} + \delta v_{(2)} - \delta \frac{dn_2}{dn_1}(v_{(1)} + (p + \tau_2^*)) = p + \tau_1^{\alpha=1}$$

where the $\tau_1^{\alpha=0}$ and $\tau_1^{\alpha=1}$ represent the optimal commodity tax when $\alpha = 0$ or 1, respectively[22].

The right-hand side of both equations can be interpreted as the marginal benefit, equal to the left-hand side, the marginal cost. The difference in the marginal utility comes from the increment of α from 0 to 1. The perceived marginal benefit for consuming n_1 changed by:

$$-\delta(u_{(2)} - v_{(2)}) + \delta \frac{dn_2}{dn_1}(u_{(1)} - v_{(1)})$$

While the first term describes the individual preventing the direct harm to the next generation, the second term describes the individual correcting the next

---

[22]See Appendix B, Equation (B.4), substitute α with value 0 or 1 to obtain the relevant result.

generation through the channel of habit. The component $\frac{dn_2}{dn_1}$ measures the elasticity of intergenerational habit formation, and $(u_{(1)} - v_{(1)})$ is the within-generation bias. Their product, therefore, represents the portion of the behavioural distortion already internalised via habits.

Efficiency Perspective:

The generation-one social planner should offset either the correction from the parents or the correction from the generation-two social planner. However, it is more efficient for the government to impose taxation than the parents to use the habit to indirectly change the next generation's behaviour.

Parents sacrifice their own welfare to re-optimise for the welfare of their child, and have a limited impact due to the $\frac{dn_2}{dn_1}$. When parents are already correcting. When the parent is already aware of the true welfare of their child, the government can improve the efficiency of the taxation by reducing the tax on the first generation to offset their own correction, and then correct the next generation directly in the next period using the optimal tax τ_2^*. Hence, the generation-one social planner reduces the τ_1 by $\delta \frac{dn_2}{dn_1}(u_{(1)} - v_{(1)})$. This justifies the result in Lemma 2.2.

Summary:

In summary, Proposition 2.2 shows that the optimal tax in the two-generation setting depends on both within-generation and inter-generation corrective terms. The IG corrective component can be decomposed into two opposing components: a positive direct correction that accounts for the external effect of parental behaviour on their children's welfare, and a negative habitual adjustment that reflects the portion of the bias already corrected by parents through habit formation.

The parameter α captures how accurately parents perceive their children's true welfare. A lower α means weaker parental correction and a higher optimal tax; a higher α means more parental self-correction, so the tax approaches the single-generation benchmark, adjusted for over-correction through intergenerational habits. Together with habit strength, α determines both the direction and size of the optimal corrective tax.

Overall, this theoretical result decomposes the optimal tax and highlights the potential over-correction and therefore loss of welfare when both the state and parents act as behavioural correctors of the next generation.

Remark:

Given the model structure, Proposition 2.2 implies that the optimal commodity tax schedule does not target the over-consumption that originates from parental

misbehaviour. If the generation-one agent over-consumes, $u(n_2, n_1)$ and $v(n_2, n_1)$ changes together. The intuition behind is that it is not the child's mistake when their unhealthy consumption stems from parental over-consumption of sugar. Hence, there is no incentive for the government to correct such inherited behaviour directly.

A common critique is: "how does the government know more about the children than the parents?" Proposition 2.2 provides a formal response to this concern. Whether the government knows more about the child's welfare than the parents is not merely philosophical but directly affects the optimal policy design. When the parents already know what is "the best for their child", the policy makers should reduce the tax rate since the correction has been done already.

2.4 Infinite Horizon Model

I extend the model to settings in which both individuals and social planners may value outcomes beyond two generations. Formally, I address this as the infinite horizon problem, where the *horizon* $T \in \mathbb{N} \cup \{\infty\}$ represents the largest number of generations that an individual or social planner can consider in an optimisation problem. Extending the optimisation problem to infinite horizon, I obtain the consumer and the social planner's problem, respectively. For a specific $t \in \mathbb{N}$, the generation-t consumer optimisation problem becomes:

$$\max_{h_t, n_t} O_t = h_t + u(n_t, n_{t-1})$$

$$+ \sum_{i=1}^{\infty} \delta^i (h_{t+i} + \alpha v(n_{t+i}, n_{t+i-1}) + (1-\alpha) u(n_{t+i}, n_{t+i-1}))$$

$$\text{s.t.} \quad h_{t+i} + (p + \tau_{t+i}) n_{t+i} = \frac{1}{1+\zeta}(z + \mathcal{T}_{t+i}) \quad \text{for all } i \geq 0$$

$$n_{t+i+1} = g(n_{t+i}), \quad \text{for all } i \geq 0$$

Suppose the optimiser of the consumer's problem is (h_t^*, n_t^*). The generation t social planner's optimisation problem becomes:

$$\max_{\tau_t, \mathcal{T}_t} W_t = h_t^* + v(n_t^*, n_{t-1}) + \sum_{i=1}^{\infty} \delta^i (h_{t+i}^* + v(n_{t+i}^*, n_{t+i-1}^*))$$

$$\text{s.t.} \quad \tau_{t+i} n_{t+i}^* = \frac{1}{1+\zeta} \mathcal{T}_{t+i} \quad i \geq 0$$

$$n_{t+i+1}^* = g(n_{t+i}^*), \quad i \geq 0$$

The consumer's objective function captures each generation's well-being and their concern for future generations, weighted by how accurately parents perceive

21

their child's true welfare (α). The social planner maximises true welfare across all generations, valuing each generation's normative utility and discounting future wellbeing. The budget constraint limits consumers' total spending after taxes and rebates, and social planners' total tax revenue equals the rebate. The habit constraint ensures that consumption patterns evolve consistently over time.

When $T = \infty$, the sequence of decisions $\{n_t\}_{t=1}^{\infty}$ may no longer depend on the finite endpoint of the problem. Instead, behaviour and policy can converge toward a long-run steady pattern that repeats across generations. To formalise this notion, I introduce the definition of *stationary*.

Definition 2.1. Let $\{x_t\}_{t=1}^{\infty}$ be a sequence of variables generated by a continuous and differentiable law of motion

$$x_{t+1} = F(x_t).$$

A point x^* is a *fixed point* of F if $F(x^*) = x^*$. The variable x_t is said to be *stationary* if, as $t \to \infty$, it converges to its fixed point:

$$\lim_{t \to \infty} x_t = x^*.$$

This definition introduces a new definition for $g(n_t)$ that induces the IG habit term $\frac{dn_{t+1}}{dn_t}$: the law of motion for the intergenerational consumption n_t. I propose the following lemma:

Lemma 2.3. *If both consumption and the optimal tax are stationary, that is,*

$$\lim_{t \to \infty} n_t = n^* \ \text{and} \ \lim_{t \to \infty} \tau_t = \tau_{\infty},$$

Then the sequence of IG habit term converges to a constant, notated as $\eta \in \mathbb{R}$:

$$\lim_{t \to \infty} \frac{dn_{t+1}}{dn_t} = g'(n^*) := \eta.$$

I have included a formal proof using the implicit function theorem in Appendix B to map the term η to the utility function. The intuition, however, is straightforward. As $t \to \infty$, the law of motion $n_{t+1} = g(n_t)$ is evaluated at the steady-state level n^*. The law of motion itself does not depend on time, as there is no endpoint in the infinite horizon problem. Hence the marginal effect $\frac{dn_{t+1}}{dn_t} = g'(n_t)$ converges to the constant $g'(n^*)$.

For the simplicity of the analysis, I assume the value of the IG habit term is independent of the value of n_t.[23] η is a parameter that characterise the changing

---

[23]At steady state, this assumption is not necessary since η is constant.

rate. Consider a simple example where n_t follows the affine law of motion: $n_{t+1} = g(n_t) = a + bn_t$, where a and b are constants. A stationary level fixed point exists when $n_{t+1} = n_t = n^*$, giving

$$n^* = \tfrac{a}{1-b}(b \neq 1) \quad and \quad g'(n^*) = b = \eta.$$

Economically, this means that the persistence of consumption behaviour across generations (η) remains constant, independent of the consumption level, reflecting stable intergenerational transmission of habits in the long run. I assume $\eta \in (0, 1)$ to obtain a well-defined and positive steady-state.

I further characterise the stationary process in consumption with the following lemma.

Lemma 2.4. *If the IG habit term is a constant η and the optimal tax is stationary, that is,*

$$\frac{dn_{t+1}}{dn_t} = \eta \text{ for all } t, \text{ and } \lim_{t\to\infty} \tau_t = \tau_\infty,$$

Then there exists a sequence of n_t that solves the optimisation problem of the consumer, and it converges to a constant, notated as n^*:

$$\lim_{t\to\infty} n_t = n^*.$$

A proof for the Lemma 2.4 is in the Appendix B, using the first order condition of the infinite horizon problem in Appendix B. The intuition is straightforward: a stationary budget constraint, affected by τ_t and a well-behaved and stationary (in fact a constant) objective function, will yield a stationary result.

I further infer the stationary property of the taxation τ, formally:

Proposition 2.3. *If $\eta \in (0, 1)$ is a constant and n_t is stationary and converge to n^* as $t \to \infty$, the optimal tax is stationary as well, moreover, it is identical to the first period optimal tax in the two generation model:*

$$\lim_{t\to\infty} \tau_t \to \tau_\infty,$$

$$and \ \tau_\infty = u_{(1)} - v_{(1)} + \delta\{(1 - \alpha)(u_{(2)} - v_{(2)}) + \alpha[-\eta(u_{(1)} - v_{(1)})]\}$$

The intuition of the proof is relatively straightforward: it is dynamically consistent across generations; therefore, the intergenerational correction for the next generation coincides with the following infinitely many generations. I have provided a formal proof in Appendix B.

2.5 Interpretation

Lemmas 2.3 and 2.4, together with Proposition 2.3, imply the existence of a stream of consumption and tax that converged to a steady state. At such a steady state, the optimal tax τ_∞ is identical to the generation one optimal tax in the two-generation model (τ_1^* in Proposition 2.2).

Intuitively, suppose the generation t is in a steady state, then so are the generations $t + 1$, $t + 2$, $t + 3$. The consumers and social planners in a steady state face the identical optimisation problem. For generation t consumers and social planners, though they do care about the welfare of $t + 2$ and the later generations, so do the generation $t + 1$ consumers and social planners. It would result in over-correcting if the generation t social planner adds another term to address the concerns over the generation $t + 2$ and later. Therefore, the optimal tax in generation t should only address the within-generation concerns and the concerns over the immediate next generation, which is effectively the identical problem of the generation-one's optimisation problem over two generations.

This equivalence highlights that, under a stationary environment and a stable habit formation process, the intergenerational optimisation problem effectively collapses to the two-generation case, and demonstrates the robustness of the two-generation result, identifies the key objective for further sufficient statistic formula derivation, and empirical identification. At the same time, this steady state requirement justifies the assumption of $\eta \in (0, 1)$.

3 Sufficient Statistics Approach

The parameters in Proposition 2.2: $u_{(1)} - v_{(1)}$ and $u_{(2)} - v_{(2)}$ are unobservable in real-world data. Hence, I adopt the sufficient statistic approach on intergenerational SSBs tax, in order to map the theoretical results above to observable data. The following propositions and simplifications are proposed to estimate the parameters, developed on the work from Allcott et al. (2019); Handel and Kolstad (2015); Bronnenberg et al. (2015).

3.1 Bridge between Theory and Data

There are two types of agents in the population: general agents and normative agents. Conceptually, a normative agent is a counterfactual of the general agent, after obtaining knowledge.[24] For each agent, their demand functions are denoted as d if they are a general agent, or d^V if they are a normative agent. General agents are biased about their own consumption, and partially biased when it comes to the next generation's consumption if they have a child, while normative agents are unbiased and behave as if they are maximising the normative utility. Specifically:

$$d(p, z, c) = \begin{cases} \text{argmax}_{n_t}(U_t), & \text{if } c = 0, \\ \text{argmax}_{n_t}(U_t + \delta(\alpha V_{t+1} + (1-\alpha)U_{t+1}), & \text{if } c = 1. \end{cases}$$

$$d^V(p, z, c) = \begin{cases} \text{argmax}_{n_t}(V_t), & \text{if } c = 0, \\ \text{argmax}_{n_t}(V_t + \delta V_{t+1}), & \text{if } c = 1. \end{cases}$$

I used a binary variable c to notate whether the agent has any children. When the agent has at least one child, I assume they maximise an intergenerational utility, with knowledge of the next generation's true utility at level α. When the agent does not have children, they instead maximise their own decision utility. Consistent with the previous notation, p is the price of SSB, and z is the income of the agent[25]. I further define $n_0 = d(p, z, 0)$, $n_1 = d(p, z, 1)$ for yhe general agents, and $n_0^V = d^V(p, z, 0)$,

---

[24]This distinction reveals a form of paternalism, as welfare is evaluated relative to what an "expert" decision-maker would choose. Following Allcott et al. (2019) this approach treats the normative agent as a benchmark, assuming that the general and normative agents are otherwise identical except for bias or information. I acknowledge this approach is not perfect, but has been widely used and is feasible in the scope of this thesis.

[25]With quasi-linear preferences, the Marshallian demand for n is independent of income z at interior solutions. At corners (when the non-negativity constraint on h binds), demand may still depend on z. I include z to facilitate later extensions on regressivity and heterogeneity, where income serves as a part of the type.

$n_1^V = d^V(p, z, 1)$ for normative agents.

To compare the differences between the general agents and the normative agents, I developed the *price metric bias* γ_0 and γ_1 based on Allcott et al. (2019). Formally:

Definition 3.1. $\gamma_c \in \mathbb{R}$ where $c \in \{0, 1\}$ is a price metric bias if:

$$d(p, z, 0) = d^V(p - \gamma_0, z - n\gamma_0, 0);$$
$$and\ d(p, z, 1) = d^V(p - \gamma_1, z - n\gamma_1, 1)$$

γ is the monetary evaluation of differences between a general and a normative agent. The general agent will behave like they are maximising the normative utility instead of the decision utility when they are faced with a different price, and such price difference is γ defined here.

Such price change by induces only substitution effect, as the income effect is offset by a reduction in $n\gamma$ at income z. The removal of the income effect in the demand function is consistent with the assumption in the previous assumption on quasi-linear preferences and the lump sum rebate of the tax revenue.

The price metric biases are important components of this thesis, as they connect both theory and the data. The rest of this section will address the linkage between γ_c and the preferences, as well as the methodology for estimating γ_c from the observable data.

3.2 Sufficient Statistics Derivation

I extended the model from Allcott et al. (2019); Handel and Kolstad (2015) into two generations; the price metric biases can be estimated from the consumption on SSBs with the following proposition:

Proposition 3.1. $\gamma_0 \approx \frac{p}{\xi_c}(\ln(n_0) - \ln(n_0^V))$, and $\gamma_1 \approx \frac{p}{\xi_c}(\ln(n_1) - \ln(n_1^V))$.

While ξ_c is the compensated price elasticity for n, and assume it is identical between the normative and general agent, and between the agents with or without a child. Furthermore, I used the following assumption when proving the proposition 3.1:

Assumption 3.1. The compensated price elasticity is identical between the normative agents and the general agents, with or without children, notated as ξ_c[26].

---

[26]With quasi-linear preference eliminating income effect, the compensated demand elasticity is equal to the (uncompensated) demand elasticity for commodity n. While c was used to notate whether one has a child, and compensated demand, I acknowledge that there is an abuse of the notation.

The proof of 3.1 is as following:

Proof. We first consider those without children ($c = 0$). Using the O notion and ignore all terms above $O(\gamma_0^2)$, I perform the first-order Taylor Expansion around (p, z):

$$\ln(n_0) - \ln(n_0^V) = \ln\left(d^V(p - \gamma_0, z - n_0\gamma_0, 0)\right) - \ln\left(n_0^V\right)$$

$$= \ln\left(n_0^V\right) - \frac{1}{n_0}\left(\frac{\partial d^V}{\partial p}\gamma_0 + \frac{\partial d^V}{\partial z}n_0\gamma_0\right) - \ln\left(n_0^V\right) + O(\gamma_0^2)$$

$$\approx -\frac{\gamma_0}{n_0}\left(\frac{\partial d^V}{\partial p} + n_0\frac{\partial d^V}{\partial z}\right)$$

The Slutsky identity: $\frac{\partial d_c^V}{\partial p} = \frac{\partial d^V}{\partial p} + n^V\frac{\partial d^V}{\partial z}$, where d_c^V is the compensated demand for the normative agent in terms of Hicksian demand. We further obtain:

$$\ln(n) - \ln(n^V) \approx -\frac{\gamma}{n^V}\left(\frac{\partial d_c^V}{\partial p}\right)$$

$$= -\frac{\gamma}{n^V}\left(\frac{\partial d_c^V}{\partial p}\frac{p}{n^V}\right)\frac{n^V}{p} = \frac{\gamma}{p}\xi_c^V$$

This allows a cleaner form of the estimation: $\ln(n_0) - \ln(n_0^V) \approx \frac{\gamma}{p}\xi_c$, therefore, $\gamma_0 \approx (\ln(n_0) - \ln(n_0^V))\frac{p}{\xi_c}$. Similarly: $\gamma_1 \approx (\ln(n_1) - \ln(n_1^V))\frac{p}{\xi_c}$ □

To link such theoretical estimation back into the formula, we propose the following theorem:

Proposition 3.2. $u_{(1)} - v_{(1)} = \gamma_0$.

The proposition 3.2 links the empirical bias term γ_0 to the theoretical marginal internality $u_{(1)} - v_{(1)}$. When the general agent consumes $\hat{n}$ facing price p, and the normative agent would choose the same $\hat{n}$ only if the price were $p - \gamma_0$, their first-order conditions imply $u_{(1)}(\hat{n}, n_{t-1}) = p$ and $v_{(1)}(\hat{n}, n_{t-1}) = p - \gamma_0$. The difference between the two marginal utilities is therefore γ_0, which represents the welfare-relevant behavioural bias determining the optimal corrective tax.

I used a similar method to identify the value of $u_{(2)} - v_{(2)}$, and I identified the following proposition.

Proposition 3.3. $u_{(2)} - v_{(2)} = \frac{1+\delta\eta}{\delta(1-\alpha)}\gamma_1 - \frac{1+(1-\alpha)\delta\eta}{\delta(1-\alpha)}\gamma_0$, when $\alpha \in [0, 1)$. $u_{(2)} - v_{(2)}$ is undefined when $\alpha = 1$.

Proof. When $\alpha \in [0, 1)$, suppose the normative agent and general agents consume at the same level and general agents are faced with the price p, then the normative

agents are faced with a price level $p - \gamma_1$. I derived the first order condition for the normative and general agents in Appendix B, which suggest:

$$p - \gamma_1 = v_{(1)} + \frac{\delta}{1 + \delta\eta} v_{(2)} \tag{3.1}$$

$$p = \frac{1}{1 + \delta\eta} \{[1 + (1 - \alpha)\delta\eta] u_{(1)} + (1 - \alpha)\delta u_{(2)} + \alpha\delta\eta v_{(1)} + \alpha\delta v_{(2)}\} \tag{3.2}$$

Subtracting Equation (3.1) from Equation (3.2) :

$$\gamma_1 = \frac{1 + \delta\eta - \alpha\delta\eta}{1 + \delta\eta} u_{(1)} + (\frac{\alpha\delta\eta}{1 + \delta\eta} - 1)v_{(1)} + \frac{(1 - \alpha)\delta}{1 + \delta\eta} u_{(2)} + \frac{\alpha\delta - \delta}{1 + \delta\eta} v_{(2)}$$

$$= \frac{1 + \delta\eta - \alpha\delta\eta}{1 + \delta\eta}(u_{(1)} - v_{(1)}) + \frac{(1 - \alpha)\delta}{1 + \delta\eta}(u_{(2)} - v_{(2)})$$

Replace $u_{(1)} - v_{(1)}$ with γ_0, after rearrangement: $u_{(2)} - v_{(2)} = \frac{1 + \delta\eta}{\delta(1 - \alpha)} \gamma_1 - \frac{1 + (1 - \alpha)\delta\eta}{\delta(1 - \alpha)} \gamma_0$

When $\alpha = 1$, the first-order condition of the general agents is suggesting:

$$p = \frac{1}{1 + \delta\eta} \{u_{(1)} + \delta\eta v_{(1)} + \delta v_{(2)}\} \tag{3.3}$$

Combine Equation (3.3) and (3.1):

$$\gamma_1 = \frac{1}{1 + \delta\eta}(u_{(1)} - v_{(1)}) = \frac{1}{1 + \delta\eta} \gamma_0$$

Since the $u_{(2)} - v_{(2)}$ does not occur, it is not defined when $\alpha = 1$. $\qquad\square$

The proposition 3.3 maps γ_0 and γ_1 from empirical observation into the implied $u_{(2)} - v_{(2)}$. The value of α comes from assumptions. For the identical γ_0, γ_1 and η, lower α implies lower $u_{(2)} - v_{(2)}$; $u_{(2)} - v_{(2)}$ approaches infinity when α approaches 1, until unidentifiable when $\alpha = 1$. This result will be clearer once the optimal tax formula is demonstrated, and I will discuss this later.

Using the Propositions 3.2 and 3.3, the formula in proposition 2.2 can be rewritten into the following cases. If $\alpha \in [0, 1)$:

$$\tau^* = \gamma_0 + \delta(1 - \alpha)\{\frac{1 + \delta\eta}{\delta(1 - \alpha)} \gamma_1 - \frac{1 + (1 - \alpha)\delta\eta}{\delta(1 - \alpha)} \gamma_0\} - \alpha\delta\eta\gamma_0$$

$$= \gamma_0 + (1 + \delta\eta)(\gamma_1 - \gamma_0) \tag{3.4}$$

Notice that α is cancelled out in the sufficient statistics expression of the optimal tax, and I will discuss this later in this section.

Define $\Delta \ln(n_c) = \ln n_c - \ln n_c^V$, where $c \in \{0, 1\}$. Using proposition 3.1, the result could be written as

$$\tau^* = \frac{p}{\xi_c} \{\Delta \ln(n_0)) + (1 + \delta\eta)(\Delta \ln(n_1) - \Delta \ln(n_0))\} \tag{3.5}$$

If $\alpha = 1$:

$$\tau^* = \gamma_0(1 - \delta\eta) \tag{3.6}$$

$$= (1 - \delta\eta)\frac{p}{\xi_c}\Delta\ln(n_0) \tag{3.7}$$

Hence why the implied $u_{(2)} - v_{(2)}$ is undefinable when α is assumed to be 1.

3.3 Discussion

I will focus the discussion around the cases where $\alpha \in [0, 1)$, and I make the following remarks.

First, The sufficient statistic formula follows the intuition closely. γ_0 and γ_1 are the bias, measured in terms of the compensated price reduction for both parents and non-parents. The first part of the sufficient statistics formula is simply γ_0, representing the bias, or internality, for those without the next generation concerns. The $\gamma_1 - \gamma_0$ is the change of bias when one transitions into parenthood. If the bias increases after having a child ($\gamma_1 > \gamma_0$), then the optimal tax with international concerns increases to address this increase in bias. Similarly, if the bias decreases after having a child ($\gamma_1 < \gamma_0$), the optimal tax with international concerns decreases as the individuals are self-correcting already. The reduction of the optimal tax prevents over-correcting. The magnitude of the increment or reduction increases when one cares more about the next generation(δ increases) or has more control over the next generation (η increases).

Second, the term $(1+\delta\eta)(\gamma_1-\gamma_0)$ captures the IG correction term in Proposition 2.2. Mechanically, this is true, since γ_0 represents $u_{(1)} - v_{(1)}$, which is the WG correction term, leaving $\gamma_1 - \gamma_0$ measures the additional motivation of the corrective tax, which is the IG correction term. When $(\gamma_1 - \gamma_0) > 0$, parents are more biased than non-parents, and the tax must increase to correct for the additional distortion that extends to their children. Conversely, if $(\gamma_1 - \gamma_0) < 0$, parents are relatively more informed and are already correcting the next generation, the government should then reduce taxation to prevent over-correcting.

Third, notice that the sufficient statistics formula of the optimal tax in Equation (3.1) shows that the optimal tax τ^* is independent of α. Beyond formalisation, the intuition behind this result follows the sufficient statistic method itself. The parameter α determines why agents deviate from the normative benchmark, while γ_0 and γ_1 capture how large those deviations are. Once these price metric biases are estimated from data, α no longer has an explicit role in the quantification process. In

contrast, δ and η cannot be absorbed into the price metric biases because they govern the intergenerational transmission of welfare and consumption habits, rather than the bias itself. They therefore remain as parameters that determine the magnitude of the optimal tax and are not absorbed in price metric biases.

Finally, the fact that τ^* is independent of α explains how the magnitude of the implied $u_{(2)} - v_{(2)}$ changes with different assumptions about α. Once the price metric biases γ_0 and γ_1 are estimated from data, the optimal tax τ^* is determined by Equation (3.4). Intuitively, consider a group of parents who are assumed to be highly informed (α close to 1), and the other are less informed (α close to 0). One can justify having the two groups have the same corrective tax by arguing that the former group has a higher $u_{(2)} - v_{(2)}$ [27]. In other words, as parents become more knowledgeable about their children's normative utility, the same observed tax must be rationalised by a larger misalignment in utility. This explained how α affects the evaluation of different corrective terms, even though it does not appear explicitly in the sufficient statistics representation of the optimal tax.

---

[27]Recall from Proposition 2.2: $\tau_1^* = u_{(1)} - v_{(1)} + \delta\{(1-\alpha)(u_{(2)} - v_{(2)}) + \alpha[-\frac{dn_2}{dn_1}(u_{(1)} - v_{(1)})]\}$.

4 Empirical Implementation

Estimating the gap in the logarithm of the SSB consumption between normative agents and the general agents, for parents and non-parents respectively, is an important part of this thesis, as the sufficient statistics γ_0 and γ_1 rely on such.

I combine the panel data from NSLY79 and NSLY97 cohorts in the National Longitudinal Survey in the U.S. to estimate such parameters. I identify whether one is a normative agent and their parental status from the data. Then I proposed a suitable cross-sectional analysis that is both achievable under the restriction of the data and approximates the causal effect of having a child on behavioural bias. Finally, I obtained a point estimate for γ_1 and γ_0, and I use this to estimate the optimal SSB tax.

4.1 Data

NSLY79 and NSLY97 are two individual-level panel datasets. Each of them follows a distinct birth cohort across time: the NLSY79 tracks individuals born between 1957 and 1964, while the NLSY97 follows those born between 1980 and 1984. I use all the waves that include questions about SSB consumption frequencies (servings per week). Specifically, I used survey responses from the NLSY79 in 2008, 2010, 2012, 2014, 2016, 2020, and 2022, and from the NLSY97 in 2009, 2010, 2011, and 2015.

Other key variables except for the SSB consumption I used include occupation data, recorded in 4-digit 2000 census code and their parental status. I use two variables to capture an individual's parental status. The first variable is whether the individual currently has a child living in the household, and is used in the regression as a time-varying indicator of whether one has concerns over the next generation. The second variable is the year when the individual's first child was born. This variable is used to verify that any observed change in SSB consumption occurs before and after the transition into parenthood, and further examine whether consumption changes as a result of introducing concerns over the next generation, rather than reflecting a pre-existing trend.

Other control variables include age, gender, race,[28] total household income, [29] family size, and an indicator for residence within a Metropolitan Statistical Area (MSA).

---

[28]Four categories available in the dataset: Black, Hispanic, mixed, others.
[29]Binned into percentiles.

4.2 Identification

Using occupation information, I identify the normative agents as medical professionals, specifically those who work in healthcare practitioner and technical occupations (occupation codes 3000 to 3540), similar to Allcott et al. (2019)'s approach, defining dietitians and nutritionists as the "normative consumers". Since I assume the medical professionals are expected to make more informed health-related consumption choices, their decisions are suitable proxies for normative decisions.[30]

Estimating γ_0 is relatively straightforward. Since the normative agents are assumed to be the counterfactual of the general agents but with a higher level of knowledge, this consumption gap can be seen as a treatment effect of increasing knowledge. Therefore, to measure the γ_0, I compare the difference in mean consumption of SSB between the normative and general agents.

$\gamma_1 - \gamma_0$ is the effect on bias from one becoming a parent. To measure such an effect, I would ideally conduct an event study to compare each individual's bias before and after becoming a parent and aggregate them into a point estimate. I did not use this method due to the lack of normative agents becoming a parent within the observation time frame.

I estimate the difference in price metric biases between parents and non-parents cross-sectionally, [31] interpreting it as the average shift in internality associated with the presence of intergenerational concern using the regression analysis. This approach does not identify within-individual behavioural changes; it provides an approximation consistent with the theoretical formulation in Equation 3.4.

More specifically, I used the following regression model.

$$\ln(n_{i,t}) = \beta_{norm}\mathcal{N}_i + \beta_{child}\mathcal{C}_{i,t} + \beta_{int}(\mathcal{N}_i \times \mathcal{C}_{i,t}) + \beta_x X_{i,t} + \mu_t + \epsilon_{i,t} \qquad (4.1)$$

where n_{it} is the consumption of SSB for agent i at time t, $\mathcal{N}_i$ is a binary variable, represent if individual i is an normative agent, $\mathcal{C}_{i,t}$ is another binary variable represent whether individual i have at least one child in the household at time t. $X_{i,t}$ represents all other demographic or household level controls, and μ_t captures time fixed effect.

I can estimate the log consumption differences using this empirical model with

---

[30]I acknowledge that this does not include behavioural bias estimation. Allcott et al. (2019) explicitly elicit bias by asking respondents to rate their agreement with statements such as "I drink soda pop or other sugar-sweetened beverages more often than I should." This type of question is not available in the NLS dataset.

[31]I acknowledge that this approach analyses the aggregated differences descriptively, and is not equivalent to the causal effect of knowledge level on consumption, or the causal effect of having a child on the price metric bias. Due to the data availability, this is the best approach available. I will discuss how to improve the internal validity of this approach later in this section.

the unbiased estimator sample mean. Specifically:

$$\Delta \ln (n_0) = \ln (n_0) - \ln (n_0^V) = -\beta_{norm}$$

$$\Delta \ln (n_1) = \ln (n_1) - \ln (n_1^V) = -(\beta_{norm} + \beta_{int}).$$

Therefore, I can write the sufficient statistics γ_0 and γ_1 in terms of regression coefficient and compensated price elasticity:

$$\gamma_0 = -\beta_{norm} \frac{p}{\xi_c} \quad \gamma_1 = -(\beta_{norm} + \beta_{int}) \frac{p}{\xi_c}.$$

Therefore, with intergenerational (IG) concern, the implied optimal tax in terms of the regression coefficients is:[32]

$$\tau_{IG}^* = -p \frac{1}{\xi_c} (\beta_{norm} + (1 + \delta\eta)\beta_{int}),$$

The optimal tax without intergenerational concern (no IG) is:[33]

$$\tau_{no\ IG}^* = u_{(1)} - v_{(1)} = \gamma_0$$

$$= -p \frac{1}{\xi_c} \beta_{norm}.$$

Improving Internal Validity

While empirically feasible, this method sacrifices internal validity due to the potential heterogeneity between those who will have a child and those who decide not to have a child. Such heterogeneity may affect SSB consumption as well.

Intuitively, the "having a child" treatment is not randomly assigned. One can argue that it is correlated with unobserved traits such as health consciousness, knowledge in nutrition, or taste in sugar, all of which can influence how much SSB a person consumes. The cross-sectional analysis groups those who do not yet have a child and those who decide they will not have a child together. As a result, the difference in sugar consumption between those with or without children may not be only due to the effect of having a child, but may also be due to those unobserved factors that affected SSB consumption. The result of the event study, which indeed estimates the causal effect on bias from having a child, is not affected by such heterogeneity, hence why it is ideal, however is not feasible due to the lack of normative agent observation as mentioned before.

This problem is especially relevant given the structure of the NLSY data. The NLSY79 and NLSY97 cohorts cover individuals at very different life stages and in different socio-economic environments. Younger cohorts may face higher public health

---

[32]Recall Proposition 3.4.

[33]Recall Proposition 2.1 and 3.2.

awareness and lower baseline SSB consumption, while more are observed as without children since they are young. Therefore, what appears as a behavioural change due to parenting could instead reflect generational differences in health awareness or cultural norms around sugar consumption.

I used the framework of average treatment effect in the style of Angrist and Pischke (2009) to formalise this issue. For each individual i, I define $\ln(n_i) = Y_i$. Consider a treatment "having a child", where $Y_i(1)$ represents the log SSB consumption if individual i becomes a parent, and $Y_i(0)$ represents such if i is not. To estimate $\gamma_1 - \gamma_0$, However, these values are not observable. I use p_i to notate if i will eventually be a parent.

$$P_i = \begin{cases} 1 & \textit{if } i \textit{ is or eventually will be a parent} \\ 0 & \textit{if } i \textit{ is not and will not be a parent} \end{cases}$$

I can observe the log SSB consumption of the parents, notated as $\{Y_i(1)|P_i = 1\}$, the log SSB consumption of those who never are parents notated as $\{Y_i(0)|P_i = 0\}$, The log SSB consumption before one becomes a parent, notated as $\{Y_i(0)|P_i = 1\}$ is under observed for the normative agent. Ideally, I would measure the average treatment effect on the treated (ATT), represented by the expected value of the change in Y_i before and after the treatment, formally:

$$\begin{aligned} ATT &= \mathbf{E}\{Y_i(1) - Y_i(0)|P_i = 1\} \\ &= \mathbf{E}\{Y_i(1)|P_i = 1\} - \mathbf{E}\{Y_i(0)|P_i = 1\}. \end{aligned}$$

In the cross-sectional analysis, I the regression coefficients pick up a different effect, namely the estimated average treatment effect on the treated, notated as $\widehat{ATT}$:

$$\widehat{ATT} = \mathbf{E}\{Y_i(1)|P_i = 1\} - \mathbf{E}\{Y_i(0)|P_i = 1 \textit{ or } P_i = 0\}.$$

The difference between is therefore:

$$ATT - \widehat{ATT} = \mathbf{E}\{Y_i(0)|P_i = 1 \textit{ or } P_i = 0\} - \mathbf{E}\{Y_i(0)|P_i = 1\}.$$

$ATT - \widehat{ATT} = 0$ if there is no individual i that satisfied $P_i = 0$, or $\mathbf{E}\{Y_i(0)|P_i = 1\} = \mathbf{E}\{Y_i(0)|P_i = 0\}$.

This suggested two ways to align $\widehat{ATT}$ with ATT. First, I add a set of control variables X, so that

$$\mathbf{E}\{Y_i(0)|P_i = 1, X = X_i\} \approx \mathbf{E}\{Y_i(0)|P_i = 0, X = X_i\}.$$

Second, I restrict the sample, where I drop all those who will not be parents, and keep only parents and potential parents in my sample. Which method to adopt

depends on the data quality, which I will discuss in the later parts.

4.3 Descriptive Results

After processing the merged NLSY79 and NLSY97 datasets, the final sample contains approximately 8000 unique individuals observed across 76000 person–year observations. Table 1 summarises the sample characteristics. The two cohorts have roughly equal gender representation, with 51% female in the NLSY79 cohort and 49% in the NLSY97 cohort, which is representative to the population. Racial composition differs slightly: 31 per cent of the NLSY79 sample identify as Black compared to 27 per cent in NLSY97, while the share of Hispanic respondents is approximately 20 per cent in both. Approximately 3 per cent of respondents are normative agents, resulting in 279 and 293 normative agents in the respective cohorts.

Variable	NLSY79	NLSY97
Observations (Time and Agent)	47,006	29,198
Number of years	7	4
Number of unique agents	8,355	8,089
Share of female	0.508	0.492
Share of Black	0.310	0.270
Share of Hispanic	0.196	0.216
Number of normative agents	279	293
Share of normative agents	0.033	0.036
Number of agents with no child observed	1,403	3,009
Number of agents with child in the household	5,162	5,431

Table 1: Sample description for NLSY79 and NLSY97

The SSB consumption differs between normative and general agents. On average, a normative agent consumes less SSB than a general agent, as expected. Table 2 reports the average weekly consumption of SSBs by agent type across the two NLSY cohorts. Overall, younger individuals in the NLSY97 cohort consume substantially more SSBs than those in the NLSY79 cohort, reflecting generational differences in dietary habits. Among general agents, average consumption rises from 1.92 servings per week in NLSY79 to 3.51 servings in NLSY97. Normative agents consume fewer SSBs compared to the normative agents in both cohorts, averaging 1.23 servings

in NLSY79 and 2.46 servings in NLSY97. This gap in consumption suggests that normative agent status is associated with lower SSB intake, supporting the assumption that normative agents provide a benchmark for bias-free consumption behaviour.

Agent type	NLSY79	NLSY97
General	1.918	3.508
Normative	1.225	2.464

Table 2: The average SSB consumption for normative and general agents, grouped by cohort (servings per week)

As mentioned previously, the event study is not available due to the data limitation, particularly the NLSY79 cohort. I have demonstrated such a limitation in Table 3. Only a single normative agent is observed before transitioning into parenthood, compared to 117 in the NLSY97 cohort. This sharp imbalance highlights the limited individual variation available for the older cohort, demonstrating the impossibility of addressing the causal estimation of "becoming a parent" on bias, using event studies. Moreover, some of the normative agents in both cohorts are either never observed with a child (84 in NLSY79 and 92 in NLSY97) or are already parents in all observed waves.

This observation motivates different strategies to address potential selection bias for different cohorts. For the NLSY79 cohort, there is only a small share of normative agents observed with consumption before transition into parenthood. I address potential heterogeneity between those who never have children and those who will eventually become parents by including a set of demographic and household controls. For example, age, income, race, family size, and area fixed effects. In contrast, for the NLSY97 cohort, where there is enough record for the consumption level before and after becoming parents, I address the potential bias by restricting the sample to those who are either current or eventual parents, excluding respondents with no evidence of having a child across all survey waves. This approach ensures that comparisons within each cohort better reflect behavioural differences associated with parental status.

Row	NLSY79	NLSY97
Observerved before having a child	1	117
Never observed with a child	84	92
Have child in household (more than 1 period)	195	201

Table 3: The count of normative agents in each cohort under different classifications

Finding the appropriate control variables is an empirical question that requires much examination and experimentation. To guide this process, I calculate the mean SSB consumption by cohort and calendar year. As demonstrated in Figure 1, the average SSB consumption exhibits both a clear downward time trend and systematic differences across cohorts, reflecting generational variation in dietary behaviour and evolving public health awareness over time. This pattern indicates the importance of including cohort and year fixed effects to account for common trends. Contrary to my prior belief, the younger cohort consumes more SSB on average compared to the older cohort. However, this is consistent with the average SSB consumption by the normative agents described in Table 2. This potentially reflects the change in taste across generations, different health conditions, or dietary requirements.
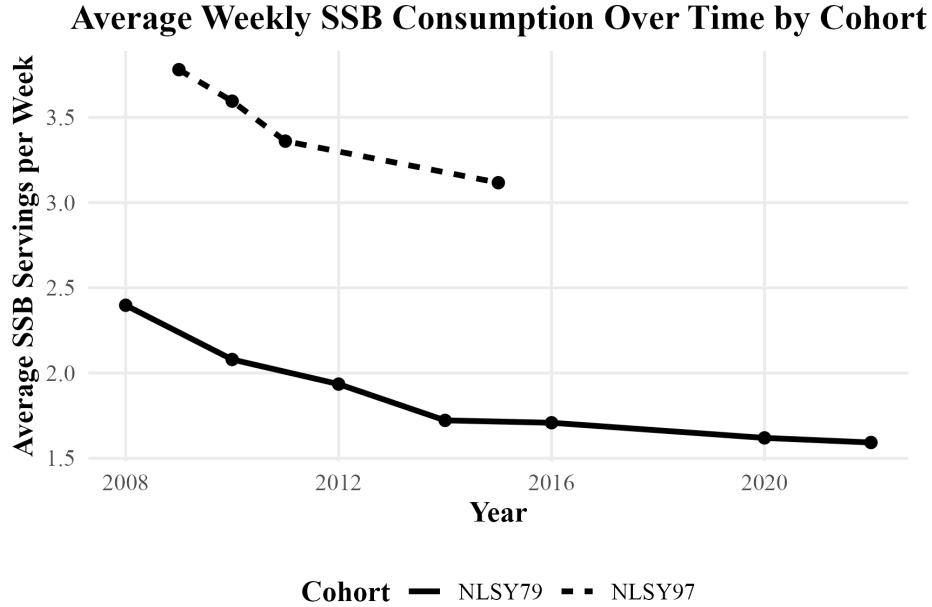

Figure 1: Average weekly SSB consumption each year, grouped by cohort

I identify individuals who will never become parents ($P_i = 0$) as NLSY79 respondents with no evidence of children in any wave. This classification is justified because by the final survey (2022), these respondents are at least 58 years old. Their SSB consumption is recorded as $(Y_i(0)|P_i = 0)$. For eventual parents ($P_i = 1$), I recover the value $(Y_i(0)|P_i = 1)$ using pre-treatment observations of SSB consumption from those who eventually become parents. This setup enables cross-sectional comparisons of the SSB consumption between those who do not yet have children and those who are never parents.

To choose the set of control variables X, so that

$$\mathbf{E}(Y_i(0)|P_i = 0, X = X_i) \approx \mathbf{E}(Y_i(0)|P_i = 1, X = X_i),$$

I proceed in four steps. First, I restrict the observations (individual at each time) to the two relevant groups: those who never become parents ($P_i = 0$) and those who eventually become parents ($P_i = 1$), but only before they become parents. Second, I construct a binary indicator I_i for parental status, where

$$I_i = \begin{cases} 1 & \textit{if } i \textit{ never have child,} \\ 0 & \textit{if } i \textit{ does not yet have child} \end{cases}$$

Third, I estimate the following regression model

$$n_{it} = \beta I_i + \beta_X X_{it} + \mu_t + \omega_c + \epsilon_{it},$$

where n_{it} denotes the weekly SSB consumption of individual i at time t, X_{it} is a vector of demographic and household-level control variables (including age, income, race, family size, and area of residence), μ_t represents time fixed effects, and ω_c represents the cohort fix effects. Finally, I sequentially expand the control set X_{it} and fixed effects to examine how the estimated coefficient β changes with different model specifications.

Table 4 presents these results, illustrating how the estimated difference in SSB consumption between eventual parents and those who never become parents evolves as additional controls are introduced.

	(1)	(2)	(3)	(4)	(5)	(6)
$\beta_{\text{indicator}}$	0.316***	0.227**	0.184**	0.146*	-0.011	0.059
	(0.085)	(0.089)	(0.088)	(0.088)	(0.086)	(0.087)
Race		✓	✓	✓	✓	✓
Income		✓	✓	✓	✓	✓
Sex			✓	✓	✓	✓
Family size				✓	✓	✓
Cohort FE	✓	✓	✓	✓		✓
Year FE	✓	✓	✓	✓	✓	

Table 4: Estimated Coefficient on Indicator (I_i) Across Alternative Control Sets

In the baseline level described in column (1), without control variables, individuals who will eventually become parents consume on average 0.32 fewer servings of SSBs per week than those who never become parents. As more variables are introduced, the difference reduces to 1.46 with the control variables: sex, race, income, and family size, as demonstrated in column (4).

By removing the year or cohort fixed effect, the SSB consumption differences between those who never have had children and those who do not yet have children quickly reduce to near 0. However, as demonstrated previously in Figure 1: both time and cohort fixed effects are important in this context, and therefore it remains empirically reasonable to include both fixed effects. One may argue that the empirical component of this thesis should prioritise minimising $\beta_{indicator}$ and therefore should drop the cohort FE. I have included the final result under this version of control (see Appendix C, Table 8 and 9), the implied optimal tax rate varies little, hence the result is robust.

To further validate the cross-sectional analysis, I estimate the dynamic relationship between SSB consumption and the timing of parenthood using an event-study framework. Specifically, I define the relative year variable as the current survey year minus the year when an individual first became a parent, captured by the year the individual received the first child, for all individuals who reported the year when they had their first child. In this construction, the individual has the relative year equal to 0 in the year they have their first child, negative values for the years before, and positive values for the years after.

This relative time structure allows me to align all individuals around the time

when they become a parent, and observe how SSB consumption evolves before and after this transition. In particular, the absence of a clear trend in SSB consumption before childbirth implies potential consumption changes are driven by the effect of becoming a parent rather than pre-existing trends.

Figure 2 plots the estimated change in SSB consumption relative to the year immediately before the birth of the first child, under alternative fixed-effect specifications. The absence of any systematic pre-trend in SSB consumption supports the validity of the identification strategy and the cross-sectional findings, which I will present later. In the years following childbirth, the estimated changes in SSB consumption remain broadly consistent across specifications during the first 10 years after parenthood, before diverging slightly in later years. This divergence highlighted the importance of controlling for appropriate fixed effects, as they influence the long-run pattern of behavioural adjustment.
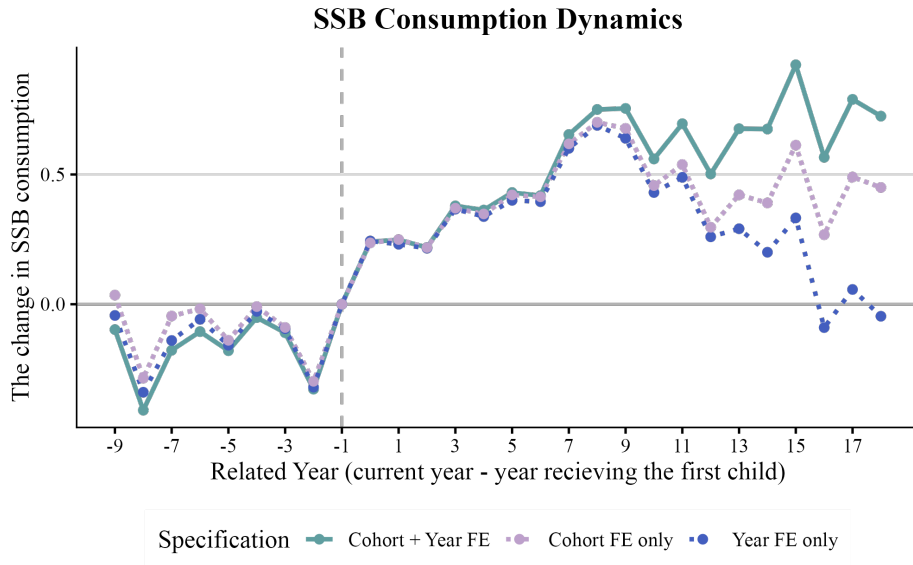

Figure 2: Change in SSB consumption by relative year to first childbirth, relative to one year before becoming a parent

In summary, the descriptive results provide strong support for the validity of the cross-sectional analysis used in this thesis, demonstrating that the dataset and empirical design lead to valid empirical estimation, thereby establishing a solid foundation for the subsequent regression analysis that formally derives the sufficient-statistic parameters for the optimal intergenerational SSB tax.

4.4 Simulating Optimal Tax

With the chosen fixed effects, control variables and the cleaned dataset, I performed the regression analysis as described in Equation (4.1). The regression result with different specifications on whether to include the discussed control variables and fixed effects is presented in Table 5. The main result is column (8), which includes the control variables, time fixed effect, and cohort fixed effect. The robustness check regarding the result of removing one fixed effect, as in columns (6) and (7), is presented in the Appendix C, Table 8 and 9.

I set the intergenerational discount factor (δ) equal to 0.99, implying that parents value their child's welfare almost as much as their own. Following Allcott et al. (2019), who estimated the average compensated price elasticity of demand for SSBs in the United States to be approximately 1.39, I adopt this value as the baseline elasticity. For comparison and robustness, I also consider two alternative scenarios with elasticities of 1 and 2. The price of SSB varies largely according to the size, brand, location and other factors. To avoid unnecessary assumptions, I simulate the optimal tax rate ($\frac{\tau*}{p}$). Therefore, the optimal tax rate for with or without intergenerational concern is:

$$\frac{\tau^*_{IG}}{p} = -\frac{1}{\xi_c}(\beta_{norm} + (1 + \delta\eta)\beta_{int})$$

$$\frac{\tau^*_{no\ IG}}{p} = -\frac{1}{\xi_c}\beta_{norm}$$

With parameters assumed, I can calculate the optimal tax rate as a function of the IG habit term η. Plotting the optimal tax rate against η, I obtain the simulation result presented in Figure 3, where the optimal tax rate ranges from 8.8% to 12.0% when the compensated price elasticity is equal to 1.39, higher than the simulated optimal SSB tax rate without intergenerational concern (no IG) 5.4%.[34]

---

[34]I acknowledge the low statistical significance of β_{int} in column (8) of Table 5, indicating that the estimated optimal tax rate with and without intergenerational concern is not significantly different. This is consistent with the simulation results and their 95% confidence intervals (see Appendix C).

Dependent Variable:				log(SSB consumption)				
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
β_{norm}	-0.175***	-0.229***	-0.214***	-0.230***	-0.0008	-0.070*	-0.055	-0.075**
	(0.037)	(0.037)	(0.036)	(0.037)	(0.038)	(0.038)	(0.038)	(0.038)
β_{child}	0.154***	0.038***	0.033***	0.021**	0.075***	0.065***	0.046***	0.055***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.013)	(0.012)	(0.012)	(0.012)
β_{int}	-0.091**	-0.022	-0.041	-0.019	-0.109**	-0.053	-0.068	-0.048
	(0.046)	(0.043)	(0.043)	(0.043)	(0.047)	(0.044)	(0.045)	(0.044)
Controls					✓	✓	✓	✓
Fixed-effects								
cohort		Yes		Yes		Yes		Yes
year			Yes	Yes			Yes	Yes
Fit statistics								
Observations	69,668	69,668	69,668	69,668	60,284	60,284	60,284	60,284
R²	0.01083	0.08907	0.08028	0.09639	0.06500	0.12848	0.11939	0.13370
RMSE	0.82389	0.79064	0.79445	0.78746	0.80197	0.77427	0.77830	0.77195

Clustered (id) standard-errors in parentheses

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Notes: This table reports the coefficient on the parental-status indicator across eight model specifications. The control variables include race, income percentile, sex, and family size. Clustered (id) standard errors are in parentheses.

Level of Significance: ***: 0.01, **: 0.05, *: 0.1

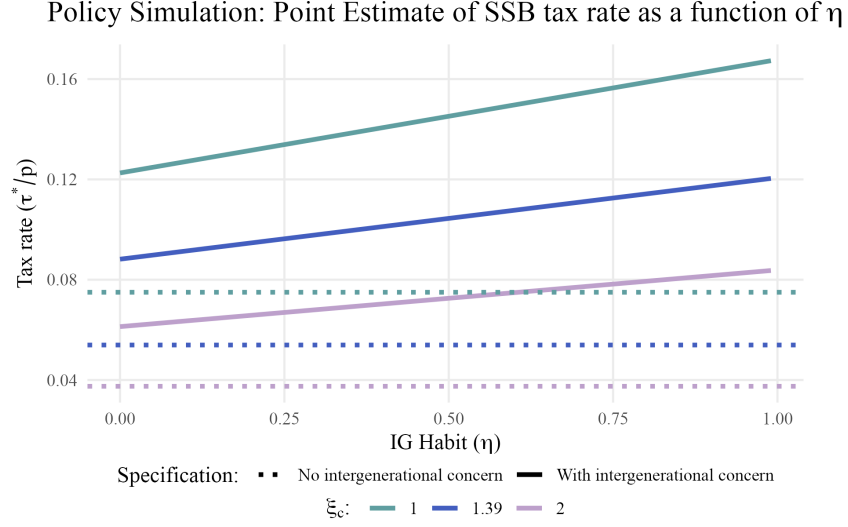**Table 5:** Main regression results for Equation (4.1)

Policy Simulation: Point Estimate of SSB tax rate as a function of η

Figure 3: Point estimate for the optimal SSB tax rate (τ^*/p) as a function of η ($\delta = 0.99$, $\xi_c \in \{1, 1.39, 2\}$)

Elasticity	no IG	$\eta = 0$	$\eta = 0.25$	$\eta = 0.5$	$\eta = 0.75$	$\eta = 0.99$
1.00	0.075	0.123	0.134	0.145	0.156	0.167
1.39	0.054	0.088	0.096	0.104	0.113	0.120
2.00	0.037	0.061	0.067	0.073	0.078	0.084

Table 6: Simulated optimal tax rate (τ^*/p) for selected elasticity and intergenerational habit term (η)

The simulation results show that incorporating intergenerational concerns increases the optimal SSB tax rate. As presented in Figure 3 and Table 6, the optimal corrective tax rises with the intergenerational habit parameter (η), reflecting that stronger parental influence on children's consumption amplifies long-term welfare losses IG direct impact, characterised in Proposition 2.2. Although the interaction term (β_{int}) is only weakly significant, the positive relationship between η and the optimal tax aligns with the theoretical discussion and sufficient statistic estimation.

From a policy perspective, the results suggest that the static internality correcting models are likely to underestimate the optimal SSB tax. Recognising this intergenerational transmission channel implies that effective tax design should account for the intergenerational impact of the parental behaviour, ensuring that corrective

taxation better reflects the long-run, multi-generational welfare implications of the SSB consumption.

5   Final Remark

The major contribution of this thesis is obtaining a point estimate of the optimal SSB tax that incorporates intergenerational concern, using measurable sufficient statistics γ_0 and γ_1. While this study focuses on SSBs, the framework could generalise to other internality correcting such as those targeting alcohol, tobacco, and excessive video games. This offers a unified framework for designing intergenerational internality corrective taxes.

This thesis develops a new framework for optimal intergenerational SSB tax and provides an empirically grounded simulation to obtain a point estimate of the welfare-maximising sugar tax. While the model is internally consistent and policy relevant, several limitations remain.

First, the theoretical framework relies on a strong homogeneity assumption, treating all agents as identical in preferences, bias magnitude, and intergenerational factors. In reality, heterogeneity across income, education, and behavioural types likely drives variation in both consumption bias and the intergenerational transmission of habits. The lack of heterogeneity across individuals limits some extensive discussion on the SSB tax, such as the regressiveness and intergenerational mobility.

Second, the empirical implementation is based on calibrated rather than causally identified relationships between variables. It is heavily limited by the data. The future research could strengthen the internal and external validity by using other quasi-experimental designs, such as estimating the treatment effect of having a child using an event study, estimating the impact of knowledge on consumption using an experiment, or estimating the intergenerational habit parameter using a natural experiment.

Finally, the choice of normative agent is a key modelling assumption. I would suggest limiting it to a nutritionist, a dietitian or those who score highly in the nutrition knowledge test for future studies. To achieve this, I would run a survey and over-sample the nutritionists and dietitians, and include a knowledge test component. This can further validate the choice for the normative agents.

44

References

Allcott, H., Lockwood, B. B. and Taubinsky, D. (2019), 'Regressive sin taxes, with an application to the optimal soda tax', *The Quarterly Journal of Economics* **134**(3), 1557–1626.
URL: *https://www.jstor.org/stable/26760744*

Allcott, H. and Taubinsky, D. (2015), 'Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market', *American Economic Review* **105**(8), 2501–38.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20131564*

Angrist, J. D. and Pischke, J.-S. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton, NJ.

Atkinson, A. B. and Stiglitz, J. E. (1976), 'The design of tax structure: Direct versus indirect taxation', *Journal of Public Economics* **6**(1), 55–75.
URL: *https://www.sciencedirect.com/science/article/pii/0047272776900414*

Attari, S. Z., DeKay, M. L., Davidson, C. I. and Bruine de Bruin, W. (2010), 'Public perceptions of energy consumption and savings', *Proceedings of the National Academy of Sciences* **107**(37), 16054–16059.
URL: *https://www.pnas.org/doi/abs/10.1073/pnas.1001509107*

Augenblick, N., Niederle, M. and Sprenger, C. (2015), 'Working over time: Dynamic inconsistency in real effort tasks *', *The Quarterly Journal of Economics* **130**(3), 1067–1115.
URL: *https://doi.org/10.1093/qje/qjv020*

Bernheim, B. D. and Rangel, A. (2009), 'Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics', *The Quarterly Journal of Economics* **124**(1), 51–104.
URL: *http://www.jstor.org/stable/40506224*

Bronnenberg, B. J., Dubé, J.-P., Gentzkow, M. and Shapiro, J. M. (2015), 'Do pharmacists buy bayer? informed shoppers and the brand premium *', *The Quarterly Journal of Economics* **130**(4), 1669–1726.
URL: *https://doi.org/10.1093/qje/qjv024*

Cawley, J. and Frisvold, D. (2023), 'Review: Taxes on sugar-sweetened beverages: Political economy, and effects on prices, purchases, and consumption', *Food Policy* **117**, 102441.
URL: *https://www.sciencedirect.com/science/article/pii/S0306919223000398*

Chetty, R. (2009), 'Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods', *Annual Review of Economics* **1**(Volume 1, 2009), 451–488.
URL: *https://doi.org/10.1146/annurev.economics.050708.142910*

Chetty, R. (2015), 'Behavioral economics and public policy: A pragmatic perspective', *American Economic Review* **105**(5), 1–33.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.p20151108*

Cheung, S. L., Tymula, A. and Wang, X. (2022), 'Present bias for monetary and dietary rewards', *Experimental Economics* **25**(4), 1202–1233.
URL: *https://doi.org/10.1007/s10683-022-09749-8*

Della Corte, K. A., Bosler, T., McClure, C., Buyken, A. E., LeCheminant, J. D., Schwingshackl, L. and Della Corte, D. (2025), 'Dietary sugar intake and incident type 2 diabetes risk: A systematic review and dose-response meta-analysis of prospective cohort studies', *Advances in Nutrition* **16**(5), 100413.
URL: *https://www.sciencedirect.com/science/article/pii/S2161831325000493*

Diamond, P. A. and Mirrlees, J. A. (1971), 'Optimal taxation and public production ii: Tax rules', *The American Economic Review* **61**(3), 261–278.
URL: *http://www.jstor.org/stable/1813425*

Dworkin, G. (1972), 'Paternalism', *The Monist* **56**(1), 64–84.
URL: *https://www.jstor.org/stable/27902250*

Gillespie, K. M., Kemps, E., White, M. J. and Bartlett, S. E. (2023), 'The impact of free sugar on human health—a narrative review', *Nutrients* **15**(4), 889.
URL: *https://www.mdpi.com/2072-6643/15/4/889*

Griffith, R., O'Connell, M. and Smith, K. (2017), 'Corrective taxation and internalities from food consumption', *CESifo Economic Studies* **64**(1), 1–14.
URL: *https://doi.org/10.1093/cesifo/ifx018*

Gruber, J. and Kőszegi, B. (2004), 'Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes', *Journal of Public Economics* **88**(9), 1959–1987.
URL: *https://www.sciencedirect.com/science/article/pii/S004727270300077X*

Handel, B. R. and Kolstad, J. T. (2015), 'Health insurance for "humans": Information frictions, plan choice, and consumer welfare', *American Economic Review* **105**(8), 2449–2500.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20131126*

Hattersley, L. and Mandeville, K. L. (2023), 'Global coverage and design of sugar-sweetened beverage taxes', *JAMA Network Open* **6**(3), e231412–e231412.
URL: *https://doi.org/10.1001/jamanetworkopen.2023.1412*

Hu, J., Megalokonomou, R. and Yuan, H. (2020), 'How do parents respond to regulation of sugary drinks in child care? evidence from california', *Journal of Economic Behavior & Organization* **178**, 672–687.
URL: *https://www.sciencedirect.com/science/article/pii/S0167268120302845*

Janzi, S., González-Padilla, E., Ramne, S., Bergwall, S., Borné, Y. and Sonestedt, E. (2024), 'Added sugar intake and its associations with incidence of seven different cardiovascular diseases in 69,705 swedish men and women', *Frontiers in Public Health* **12**.
URL: *https://doi.org/10.3389/fpubh.2024.1452085*

Lockwood, B. B. (2020), 'Optimal income taxation with present bias', *American Economic Journal: Economic Policy* **12**(4), 298–327.
URL: *https://www.aeaweb.org/articles?id=10.1257/pol.20180762*

Mill, J. S. (1859), *On Liberty*, J.W. Parker and son.
URL: *https://books.google.com.au/books?id=AjpGAAAAcAAJ*

O'Donoghue, T. and Rabin, M. (2006), 'Optimal sin taxes', *Journal of Public Economics* **90**(10), 1825–1849.
URL: *https://www.sciencedirect.com/science/article/pii/S0047272706000247*

Onaolapo, A. Y., Onaolapo, O. J. and Olowe, O. A. (2020), *Chapter 9 - An overview of addiction to sugar*, Academic Press, pp. 195–216.
URL: *https://www.sciencedirect.com/science/article/pii/B9780128169186000093*

Paraje, R. G., Jha, P., Savedoff, W. and Fuchs, A. (2023), 'Taxation of tobacco, alcohol, and sugar-sweetened beverages: reviewing the evidence and dispelling the myths', *BMJ Global Health* **8**(Suppl 8), e011866.
URL: *https://ghsite-bmj.vercel.app/content/8/Suppl_8/e011866*

Parliamentary Budget Office (2024), Policy costing: 20% tax on Sugar-Sweetened Beverages, Technical Report PR-2024-880, Parliamentary Budget Office, Parliament of Australia, Canberra, Australia. Costing completed 27 March 2024; Confidential – Authorised for public release on 24 June 2024.

Pigou, A. (1920), *The economics of welfare*, Routledge.

Rhodes, K., Chan, F., Prichard, I., Coveney, J., Ward, P. and Wilson, C. (2016), 'Intergenerational transmission of dietary behaviours: A qualitative study of anglo-australian, chinese-australian and italian-australian three-generation families', *Appetite* **103**, 309–317.
URL: *https://www.sciencedirect.com/science/article/pii/S0195666316301659*

Sadoff, S., Samek, A. and Sprenger, C. (2019), 'Dynamic inconsistency in food choice: Experimental evidence from two food deserts', *The Review of Economic Studies* **87**(4), 1954–1988.
URL: *https://doi.org/10.1093/restud/rdz030*

Saez, E. (2001), 'Using elasticities to derive optimal income tax rates', *Review of Economic Studies* **68**, 205–29.

Sharma, A., Hauck, K., Hollingsworth, B. and Siciliani, L. (2014), 'The effects of taxing sugar-sweetened beverages across different income groups', *Health Econ* **23**(9), 1159–84.

Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, number smith1776 *in* 'History of Economic Thought Books', McMaster University Archive for the History of Economic Thought.
URL: *https://econpapers.repec.org/bookchap/hayhetboo/smith1776.htm*

Talati, Z., Charlesworth, J., Mackenzie, K., McAlpine, T., Myers, G., Miller, C., Dana, L. M., O'Connor, M., Mullan, B. A. and Dixon, H. G. (2024), 'Exploring sugary drink consumption and perceptions among primary-school-aged children and parents in australia', *Nutrients* **16**(19).

U.S. Bureau of Labor Statistics (2024*a*), 'National longitudinal survey of youth 1979 (nlsy79)'. Washington, D.C. Accessed via the NLS Investigator, https://www.nlsinfo.org/investigator/.

U.S. Bureau of Labor Statistics (2024*b*), 'National longitudinal survey of youth 1997 (nlsy97)'. Washington, D.C. Accessed via the NLS Investigator, https://www.nlsinfo.org/investigator/.

Westwater, M. L., Fletcher, P. C. and Ziauddeen, H. (2016), 'Sugar addiction: the state of the science', *European Journal of Nutrition* **55**(S2), 55–69.
URL: *http://dx.doi.org/10.1007/s00394-016-1229-6*

Westwater, M. L., Fletcher, P. C. and Ziauddeen, H. (2018), 'Sugar addiction: From evolution to revolution', *Frontiers in Psychiatry* **9**, 545. Section: Psychological Therapy and Psychosomatics. Part of the Research Topic: *Food and Its Effect on the Brain: From Physiological to Compulsive Consumption.*
URL: *https://doi.org/10.3389/fpsyt.2018.00545*

Zhen, C., Wohlgenant, M. K., Karns, S. and Kaufman, P. (2011), 'Habit formation and demand for sugar-sweetened beverages', *American Journal of Agricultural Economics* **93**(1), 175–193.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1093/ajae/aaq155*

A   Table of Variables

Symbol	Definition / Description
p	Price of sugar-sweetened beverages (SSB)
τ_t	SSB tax at time t
τ^*	Optimal SSB tax
τ_∞	SSB tax at steady-state levels
T_t	Lump-sum transfer or rebate from tax revenue at time t
h_t	Composite good consumption (numeraire)
n_t	Quantity of SSB consumed by generation t
z	Income or endowment of the individual
U_t	Decision utility function (biased or perceived utility)
V_t	Normative utility function (true welfare utility)
$u(n_t, n_{t-1})$	utility component from SSB consumption $(h_t + u(n_t, n_{t-1}))$
$v(n_t, n_{t-1})$	Normative (true) utility component from SSB consumption
$u_{(1)}$	Partial derivatives of $u(\cdot)$ w.r.t. first arguments
$u_{(2)}$	Partial derivatives of $u(\cdot)$ w.r.t. second arguments
$v_{(1)}, v_{(2)}$	Similar to $u_{(1)}, u_{(2)}$
ζ	Share of parental consumption passed to the child
δ	Intergenerational discount factor (parental concern for child's welfare)
α	Accuracy of parental perception of child's true welfare
η	Intergenerational habit transmission parameter
$g(n_t)$	Law of motion for intergenerational consumption
$g'(n^*) = \eta$	Derivative of law of motion at steady state (defines η)
n^*	Steady-state SSB consumption across generations
$d(p, z, c)$	Demand function of a general agent
$d^V(p, z, c)$	Demand function of a normative (unbiased) agent
n_0, n_1	SSB consumption by non-parents and parents (general agents)
n_0^V, n_1^V	SSB consumption by normative non-parents and normative parents
γ_0, γ_1	Price metric bias for non-parents and parents (measured in \$)
ξ_c	Compensated price elasticity of demand for SSB

Table 7: List of Variables and Parameters

B Proofs

Normative and Decision Utility Framework

In the traditional dynamic inconsistency model, individuals maximise

$$U_t = u(c_t) + \beta \sum_{s=t+1}^{\infty} \delta^{s-t} u(c_s), \; \delta, \beta \in [0,1].$$

Where t is a time within one's lifetime, and c_t is the consumption at period t, $u(c_t)$ is the instantaneous utility in period t, δ is the exponential discounting factor and β captures present bias. Hence, I can interpret U_t as one's decision utility for a lifetime observed at time t.

The normative benchmark in this case is instead:

$$V_t = u(c_t) + \sum_{s=t+1}^{\infty} \delta^{s-t} u(c_s), \; , \; \delta \in [0,1].$$

Similarly, I can interpret the V_t as one's normative utility for a lifetime observed at time t. Therefore, one can interpret the dynamic inconsistency as a special case of normative and decision utility, that provides a mechanism for present bias using parameter β.

The proof for Proposition 2.2

Proof. We extend the single-generation argument by allowing the parents' consumption to affect the child's via $n_2 = g(n_1)$ and by letting parents weight the next generation's *normative* utility with knowledge parameter $\alpha \in [0,1]$ and discount factor $\delta \in (0,1)$. I prove the Proposition 2.2 in the following 3 steps:

1. **The social planner chooses the optimal consumption:**

   The social planner's problem is

   $$\max_{h_1^*, n_1^*, h_2^*, n_2^*} W_1 = h_1^* + v(n_1^*, n_0) + \delta[h_2^* + v(n_2^*, n_1^*)]$$

   subject to feasibility and habit formation:

   $$h_1^* + (p + \tau_1) n_1^* = \frac{1}{1 + \zeta}(z + \mathcal{T}_1) \tag{B.1}$$

   $$\mathcal{T}_1 = (1 + \zeta) n_1^* \tau_1 \tag{B.2}$$

   $$h_2 + (p + \tau_2) n_2 = z + \mathcal{T}_2$$

   $$n_2 = g(n_1)$$

   Notice that $\tau_2 = \tau_2^*$ *and* $\mathcal{T}_2 = \mathcal{T}_2^*$ suggested $h_2 + p n_2 = z$, and Equation (B.1) and (B.2) suggested:

   $$h_1^* + p n_1^* = z$$

51

Let multipliers $(\lambda_1, \lambda_2, \theta)$ correspond to these three constraints. The Lagrangian is

$$\mathcal{L} = h_1 + v(n_1, n_0) + \delta[h_2 + v(n_2, n_1)] - \lambda_1(h_1 + pn_1 - z)$$
$$- \lambda_2(h_2 + pn_2 - z)$$
$$- \theta[n_2 - g(n_1)]$$

The first order condition suggested:

$$\frac{\partial \mathcal{L}}{\partial h_1} = 1 - \lambda_1 = 0 \quad \frac{\partial \mathcal{L}}{\partial h_2} = \delta - \lambda_2 = 0$$
$$\frac{\partial \mathcal{L}}{\partial n_1} = \delta v_{(1)} - \delta p - \theta = 0 \quad \frac{\partial \mathcal{L}}{\partial n_2} = v_{(1)} + \delta v_{(2)} - p + \theta \frac{dn_2}{dn_1} = 0$$

in h_1, h_2 yield $\lambda_1 = 1$ and $\lambda_2 = \delta$. The FOCs in n_2 and n_1 give

$$\theta = \delta(v_{(1)} - p) \; and \; v_{(1)} + \delta v_{(2)} - p + \theta \frac{dn_2}{dn_1} = 0.$$

Substituting θ,

$$p = v_{(1)} + \delta v_{(2)} + \delta \frac{dn_2}{dn_1}(v_{(1)} - p) \tag{B.3}$$

2. **Consumers choose according to τ_1**

   The consumers (the parents) in generation one are faced with the following consumer problem

   $$\max_{h_1, n_1, h_2, n_2} O_1 = h_1 + u(n_1, n_0) + \delta[h_2 + \alpha v(n_2, n_1) + (1 - \alpha)u(n_2, n_1)]$$

   subject to

   $$h_1 + (p + \tau_1)n_1 = \tfrac{1}{1+\zeta}(z + \mathcal{T}_1)$$
   $$h_2 + (p + \tau_2)n_2 = z + \mathcal{T}_2$$
   $$n_2 = g(n_1)$$

   where $\mathcal{T}_t$ is lump-sum (independent of individual n_t). Let multipliers $(\lambda_1, \lambda_2, \theta)$ correspond to these constraints. The Lagrangian is

   $$\mathcal{L} = h_1 + u(n_1, n_0) + \delta[h_2 + \alpha v(n_2, n_1) + (1 - \alpha)u(n_2, n_1)]$$
   $$- \lambda_1(h_1 + (p + \tau_1)n_1 - \tfrac{1}{1+\zeta}(z + \mathcal{T}_1))$$
   $$- \lambda_2(h_2 + (p + \tau_2)n_2 - (z + \mathcal{T}_2))$$
   $$- \theta[n_2 - g(n_1)]$$

The first order condition suggested:

$$\frac{\partial \mathcal{L}}{\partial h_1} = 1 - \lambda_1 = 0$$

$$\frac{\partial \mathcal{L}}{\partial h_2} = \delta - \lambda_2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_1} = u_{(1)} - (p + \tau_1) + \delta(\alpha v_{(2)} + (1 - \alpha)u_{(2)}) - \theta\frac{dn_2}{dn_1} = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_2} = \delta(\alpha v_{(1)} + (1 - \alpha)u_{(1)}) - \delta(p + \tau_2) - \theta = 0$$

This imply $\lambda_1 = 1$ and $\lambda_2 = \delta$. The first order condition in n_2 yields

$$\theta = \delta(\alpha v_{(1)} + (1 - \alpha)u_{(1)} - (p + \tau_2)),$$

Substitute θ into $\frac{\partial \mathcal{L}}{\partial n_1}$

$$p + \tau_1 = u_{(1)} + \delta(\alpha v_{(2)} + (1-\alpha)u_{(2)}) - \delta\frac{dn_2}{dn_1}(\alpha v_{(1)} + (1-\alpha)u_{(1)} - (p + \tau_2)) \text{ (B.4)}$$

3. **Aligning Social Planner's choice and Individual rationality**

   For generation 2, comparing the planner's n_2 condition with the child's consumer condition (the usual single-generation step) gives

   $$\tau_2^* = u_{(1)} - v_{(1)}$$

   To have the (h_1^*, n_1^*) picked by the social planner chosen by the individual and satisfy the individual rationality, Equation (B.3) and (B.4) both hold. I then insert τ_2^* to eliminate p. After rearrangement,

   $$\tau_1^* = (u_{(1)} - v_{(1)}) + \delta[(1 - \alpha)(u_{(2)} - v_{(2)}) - \alpha\frac{dn_2}{dn_1}(u_{(1)} - v_{(1)})].$$

Therefore,

$$\tau_1^* = (u_{(1)} - v_{(1)}) + \delta[(1 - \alpha)(u_{(2)} - v_{(2)}) - \alpha g'(n_1)(u_{(1)} - v_{(1)})]$$
$$\tau_2^* = u_{(1)} - v_{(1)}$$

which proves Proposition 2.2. □

Details of the proof for proposition 3.3

For the normative agent, the optimisation problem is

$$\max_{h_1, n_1, h_2, n_2} W = h_1 + v(n_1, n_0) + \delta(h_2 + v(n_2, n_1))$$

53

subject to $h_1 + (p - \gamma_1)n_1 = z$, $h_2 + (p - \gamma_1)n_2 = \bar{z}$, $n_2 = g(n_1)$, where γ_1 denotes the bias metric capturing the internality wedge between perceived and normative prices, and $g'(n_1) = \eta$ represents the intergenerational habit elasticity, $\bar{z} = \frac{1}{1+\zeta}z$. The Lagrangian is

$$\mathcal{L} = h_1 + v(n_1, n_0) + \delta(h_2 + v(n_2, n_1))$$
$$- \lambda[h_1 + (p - \gamma_1)n_1 - \bar{z}] - \mu[h_2 + (p - \gamma_1)n_2 - \bar{z}] - \theta[n_2 - g(n_1)].$$

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial h_1} = 1 - \lambda = 0, \qquad\qquad \frac{\partial \mathcal{L}}{\partial h_2} = \delta - \mu = 0,$$

$$\frac{\partial \mathcal{L}}{\partial n_1} = v_{(1)} + \delta v_{(2)} - \lambda(p - \gamma_1) + \theta g'(n_1) = 0, \quad \frac{\partial \mathcal{L}}{\partial n_2} = \delta v_{(1)} - \mu(p - \gamma_1) - \theta = 0.$$

Substituting $\lambda = 1$, $\mu = \delta$, and $g'(n_1) = \eta$:

$$v_{(1)} + \delta v_{(2)} - (p - \gamma_1) + \eta\theta = 0, \qquad \delta v_{(1)} - \delta(p - \gamma_1) - \theta = 0.$$

Solving yields $\theta = \delta[v_{(1)} - (p - \gamma_1)]$, and substituting back:

$$v_{(1)} + \delta v_{(2)} - (p - \gamma_1) + \eta\delta[v_{(1)} - (p - \gamma_1)] = 0.$$

After rearrange, the result is as following;

$$\boxed{p - \gamma_1 = v_{(1)} + \frac{\delta}{1 + \delta\eta}v_{(2)}.}$$

For general agents, the optimisation problem is:

$$\max_{h_1, n_1, h_2, n_2} O = h_1 + u(n_1, n_0) + \delta(h_2 + \alpha v(n_2, n_1) + (1 - \alpha)u(n_2, n_1))$$

subject to $h_1 + pn_1 = z$, $h_2 + pn_2 = \bar{z}$, $n_2 = g(n_1)$, where $g'(n_1) = \eta$ represents the intergenerational habit elasticity, $\bar{z} = \frac{1}{1+\zeta}z$. The Lagrangian is

$$\mathcal{L} = h_1 + u(n_1, n_0) + \delta(h_2 + \alpha v(n_2, n_1) + (1 - \alpha)u(n_2, n_1))$$
$$- \lambda[h_1 + pn_1 - \bar{z}] - \mu[h_2 + pn_2 - \bar{z}] - \theta[n_2 - g(n_1)].$$

The Lagrangian FOCs are

$$\frac{\partial \mathcal{L}}{\partial h_1} = 1 - \lambda = 0 \ and \ \frac{\partial \mathcal{L}}{\partial h_2} = \delta - \mu = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_1} = u_{(1)} + \delta(\alpha v_{(2)} + (1 - \alpha)u_{(2)}) - \lambda p + \theta\eta = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_2} = \delta(\alpha v_{(1)} + (1 - \alpha)u_{(1)}) - \mu p - \theta = 0$$

With $\lambda = 1, \mu = \delta$, we get $\theta = \delta(\alpha v_{(1)} + (1 - \alpha)u_{(1)} - p)$, hence $u_{(1)} + \delta(\alpha v_{(2)} + (1 - \alpha)u_{(2)}) - p + \eta\delta(\alpha v_{(1)} + (1 - \alpha)u_{(1)} - p) = 0$. After rearrange, the result is as following:

$$p = \frac{1}{1 + \delta\eta}\{[1 + (1 - \alpha)\delta\eta]u_{(1)} + (1 - \alpha)\delta u_{(2)} + \alpha\delta\eta v_{(1)} + \alpha\delta v_{(2)}\}$$

The Proof for Lemma 2.3

Proof. For those in period t, their objective function is:

$$O_t = U_t + \sum_{i=t+1}^{\infty} \delta(\alpha V_i + (1 - \alpha)U_i)$$

Notice that for each case of knowledge, the objective function is time-invariant. Formally; $O_t = O_j$ for all t, j. The first order condition suggests:

$$\frac{\partial O_t}{\partial n_t}(n_t^*, n_{t-1}) = p + \tau_\infty$$

We define this as: $O_{(1)}(n_t^*, n_{t-1}) = p + \tau_\infty$

This result is invariant regarding n_{t-1}, hence we can write out:

$$\frac{\partial O_{(1)}}{\partial n_{t-1}}(n_t^*, n_{t-1}) = 0 \tag{B.5}$$

We can view the optimal consumption level in the second period as a function of the consumption level in the first period. The Equation B.5 can be we written as:

$$\frac{\partial O_{(1)}}{\partial n_{t-1}}(n_t^*(n_{t-1}), n_{t-1}) = 0$$

$$\frac{dn_t^*}{dn_{t-1}}\frac{\partial O_{(1)}}{\partial n_t^*(n_{t-1})} + \frac{\partial O_{(1)}}{\partial n_{t-1}} = \frac{dn_t^*}{dn_{t-1}}O_{(11)} + O_{(12)} = 0$$

$$\frac{dn_t^*}{dn_{t-1}} = -\frac{O_{(12)}}{O_{(11)}} \tag{B.6}$$

Given that consumption is stationary, O is time invariant, the direct impact of the parental consumption $\frac{dn_t^*}{d_{t-1}}$ is a constant, therefore stationary. $\square$

The Infinite Horizon FoC (For Lemma 2.4 and Proposition2.3)

Recall the generation-t consumer's problem:

$$\max_{h_t, n_t} O_t = h_t + u(n_t, n_{t-1})$$

$$+ \sum_{i=1}^{\infty} \delta^i (h_{t+i} + \alpha v(n_{t+i}, n_{t+i-1}) + (1-\alpha)u(n_{t+i}, n_{t+i-1}))$$

$$\text{s.t.} \quad h_{t+i} + (p + \tau_{t+i})n_{t+i} = \frac{1}{1+\zeta}(z + \mathcal{T}_{t+i}) \quad \text{for all } i \geq 0$$

$$n_{t+i+1} = g(n_{t+i}), \quad \text{for all } i \geq 0$$

I can obtain the following Lagrangian:

$$\mathcal{L} = h_t + u(n_t, n_{t-1}) + \sum_{i=1}^{\infty} \delta^i \Big(h_{t+i} + \alpha v(n_{t+i}, n_{t+i-1}) + (1-\alpha)u(n_{t+i}, n_{t+i-1}) \Big)$$

$$- \sum_{i=0}^{\infty} \lambda_i \Big(h_{t+i} + (p + \tau_{t+i})n_{t+i} - \frac{z + \mathcal{T}_{t+i}}{1+\zeta} \Big) - \sum_{i=0}^{\infty} \mu_i \Big(n_{t+i+1} - g(n_{t+i}) \Big).$$

I can then obtain the partial derivative:

$$\frac{\partial L}{\partial h_{t+i}} = 0 = \delta^i - \lambda_i \quad \text{for all } i \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial n_t} = u_{(1)} + \delta[\alpha\, v_{(2)} + (1-\alpha)\, u_{(2)}] - \lambda_0\,(p + \tau_t) + \mu_0\, g'(n_t) = 0$$

$$\frac{\partial \mathcal{L}}{\partial n_{t+i}} = \delta^i[\alpha\, v_{(1)} + (1-\alpha)\, u_{(1)}] + \delta^{i+1}[\alpha\, v_{(2)} + (1-\alpha)\, u_{(2)}]$$

$$- \lambda_i\,(p + \tau_{t+i}) + \mu_i\, g'(n_{t+i}) - \mu_{i-1} = 0 \quad \text{for all } i \geq 1$$

Notice: $g'(n_{t+i}) = \eta$. Therefore:

$$\delta^i = \lambda_i \quad \text{for all } i \geq 0$$

$$\mu_0 \eta = -u_{(1)} - \delta(\alpha v_{(2)} + (1-\alpha)u_{(2)}) + (p + \tau_t)$$

$$\mu_i \eta = \mu_{i-1} - \delta^i \underbrace{(\alpha\, v_{(1)} + (1-\alpha)\, u_{(1)})}_{\text{MB at } t+i}$$

$$+ \delta \underbrace{[\alpha\, v_{(2)} + (1-\alpha)\, u_{(2)}]}_{\text{MB at } t+i+1} - \underbrace{(p + \tau_{t+i}))}_{MC} \quad \text{for all } i \geq 1$$

The expression about the multiplier μ_i times η depends on the previous multiplier μ_i, and minus the period $t+i$ consumer problem perceived at t: the marginal benefit for the $t+1$ generation, plus the discounted marginal benefit for the $t+1$ generation observed by the generation t observer, minus the marginal cost (post tax price).

Therefore, I can construct the Euler equation using the first-order condition:

$$p + \tau_t = u_{(1)} + \delta[\alpha v_{(2)} + (1 - \alpha)u_{(2)}] + \eta\mu_0 \tag{B.7}$$

$$p + \tau_{t+i} = \alpha\, v_{(1)} + (1 - \alpha)\, u_{(1)} + \delta\left[\alpha\, v_{(2)} + (1 - \alpha)\, u_{(2)}\right]$$
$$+ \frac{\eta\,\mu_i - \mu_{i-1}}{\delta^i}, \quad i \geq 1 \tag{B.8}$$

Similarly, recall the generation-t social planner's problem:

$$\max_{\tau_t, \mathcal{T}_t} W_t = h_t^* + v(n_t^*, n_{t-1}) + \sum_{i=1}^{\infty} \delta^i(h_{t+i}^* + v(n_{t+i}^*, n_{t+i-1}^*))$$

$$\text{s.t.} \quad \tau_{t+i} n_{t+i}^* = \frac{1}{1 + \zeta}\mathcal{T}_{t+i} \quad i \geq 0$$

$$n_{t+i+1}^* = g(n_{t+i}^*), \quad i \geq 0$$

Due to the fiscal feasibility (rebate equal to the tax revenue), the budget constraint for each generation about h_x^* and n_x^* is $h_x^* + pn_x^* = \frac{z}{1+\zeta}$

I can write out the Lagrangian for the social planner:

$$\mathcal{L} = h_t^* + v(n_t^*, n_{t-1}^*) + \sum_{i=1}^{\infty} \delta^i\left(h_{t+i}^* + v(n_{t+i}^*, n_{t+i-1}^*)\right)$$

$$- \sum_{i=0}^{\infty} \tilde{\lambda}_i(h_{t+i}^* + pn_{t+i}^* - \frac{z}{1+\zeta}) - \sum_{i=0}^{\infty} \tilde{\mu}_i\left(n_{t+i+1}^* - g(n_{t+i}^*)\right).$$

I can construct the Euler equation symmetry to the consumer's problem:

$$p = v_{(1)} + \delta v_{(2)} + \eta\tilde{\mu}_0 \tag{B.9}$$

$$p = v_{(1)} + \delta v_{(2)} + \frac{\eta\,\tilde{\mu}_i - \tilde{\mu}_{i-1}}{\delta^i}, \quad i \geq 1 \tag{B.10}$$

The Proof for Lemma 2.4

Proof. When η is a constant and the tax converges to τ_∞, it is easy to see that there exist a certain n^*, such that this n^* yields a specific $u_{(1)}^*, u_{(2)}^*, v_{(1)}^*, v_{(2)}^*$, and the following condition holds:

$$p + \tau_\infty = \alpha\, v_{(1)}^* + (1 - \alpha)\, u_{(1)}^* + \delta\left[\alpha\, v_{(2)}^* + (1 - \alpha)\, u_{(2)}^*\right]$$

Hence, there exists a stationary consumption that converges to n^*. $\qquad\square$

The Proof for proposition 2.3

Proof. Consider the first-order conditions in the Appendix B, Equation B.7, B.8, B.9, B.10: For the consumers:

$$p + \tau_t = u_{(1)} + \delta[\alpha v_{(2)} + (1 - \alpha)u_{(2)}] + \eta\mu_0$$
$$p + \tau_{t+i} = \alpha v_{(1)} + (1 - \alpha)u_{(1)} + \delta\left[\alpha v_{(2)} + (1 - \alpha)u_{(2)}\right] + \frac{\eta\mu_i - \mu_{i-1}}{\delta^i}, \quad i \geq 1$$

And for the planners:

$$p = v_{(1)} + \delta v_{(2)} + \eta\tilde{\mu}_0$$
$$p = v_{(1)} + \delta v_{(2)} + \frac{\eta\tilde{\mu}_i - \tilde{\mu_{i-1}}}{\delta^i}, \quad i \geq 1$$

where $\mu_i, \tilde{\mu}_i$ are the Lagrangian multipliers, for the consumers and social planners, respectively.

I define the terms: A_i, B_i

$$A_i := \alpha v_{(1)} + (1 - \alpha)u_{(1)} + \delta\left[\alpha v_{(2)} + (1 - \alpha)u_{(2)}\right], \qquad B_i := v_{(1)} + \delta v_{(2)}.$$

The consumer and the social planner's FOC yields:

$$p + \tau_t = A_0 + \eta\mu_0,$$
$$p = B_0 + \eta\tilde{\mu}_0,$$

And for $i \geq 1$

$$p + \tau_{t+i} = A_i + \frac{\eta\mu_i - \mu_{i-1}}{\delta^i}, \quad i \geq 1$$
$$p = B_i + \frac{\eta\tilde{\mu}_i - \tilde{\mu_{i-1}}}{\delta^i}, \quad i \geq 1$$

where $\mu_i, \tilde{\mu}_i$ are the multipliers on the habit constraint.

Define $\Delta_i := \mu_i - \tilde{\mu}_i$. Subtract the planner's FOCs from the consumer's FOCs:

$$\tau_t = (A_0 - B_0) + \eta\Delta_0, \tag{B.11}$$
$$\tau_{t+i} = (A_i - B_i) + \frac{\eta\Delta_i - \Delta_{i-1}}{\delta^i}, \quad i \geq 1. \tag{B.12}$$

Assume a stationary consumption exists: $n_{t+i} \to n^*$, $g'(n^*) = \eta$, and $\delta\eta < 1$ since $\delta, \eta \in (0, 1)$. Then $A_i \to A$, $B_i \to B$, and $\tau_{t+i} \to \tau_\infty$.

Evaluating (B.12) at i and $i + 1$ and subtracting shows that $(\eta\Delta_i - \Delta_{i-1})/\delta^i$ is constant; call it C. Hence from (B.11) and (B.12):

$$\tau_\infty = A - B + C, \qquad \tau_\infty = A - B + \eta\Delta_0 \implies C = \eta\Delta_0. \tag{B.13}$$

At the fixed point,

$$A - B = (1 - \alpha)\Big[(u_{(1)} - v_{(1)}) + \delta(u_{(2)} - v_{(2)})\Big]. \tag{B.14}$$

To pin down Δ_0, use the steady-state co-state recursion implied by the Lagrangian (consumer minus planner at $i = 0$):

$$\eta\Delta_0 = \tau_\infty - \Big[(u_{(1)} - v_{(1)}) + \delta(1 - \alpha)(u_{(2)} - v_{(2)})\Big]. \tag{B.15}$$

Combine (B.13)–(B.15) and solve for τ_∞:

$$\tau_\infty = (A - B) + \eta\Delta_0$$
$$= (1 - \alpha)\Big[(u_{(1)} - v_{(1)}) + \delta(u_{(2)} - v_{(2)})\Big] + \eta\Big[\tau_\infty - (u_{(1)} - v_{(1)}) - \delta(1 - \alpha)(u_{(2)} - v_{(2)})\Big].$$

Rearranging gives

$$\tau_\infty = (u_{(1)} - v_{(1)}) + \delta\Big[(1 - \alpha)(u_{(2)} - v_{(2)}) - \alpha\eta(u_{(1)} - v_{(1)})\Big]$$

where all marginal terms are evaluated at (n^*, n^*). $\qquad\qquad\square$

C   Additional Empirical Discussion

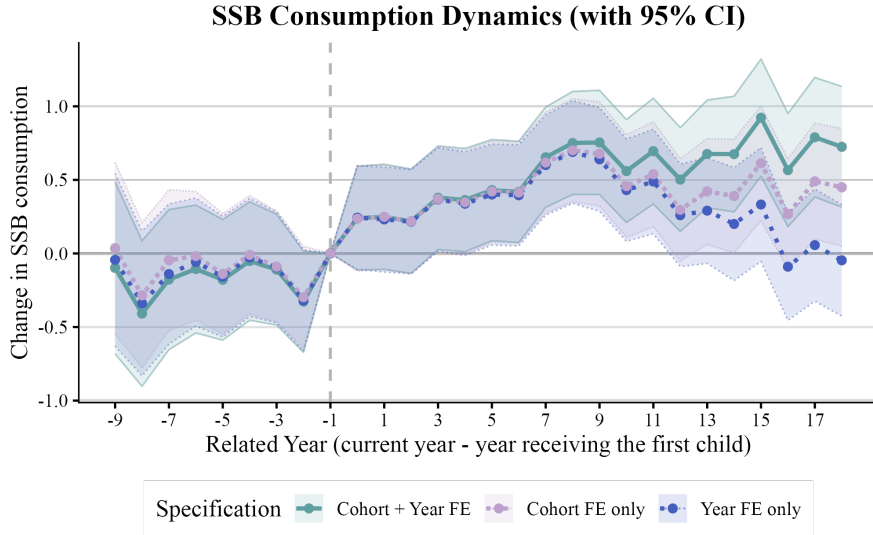## Impact on SSB from the Relative Year: CI for Figure 2

Figure 4: Change in SSB consumption by relative year to first childbirth, relative to one year before becoming a parent, 95% CI
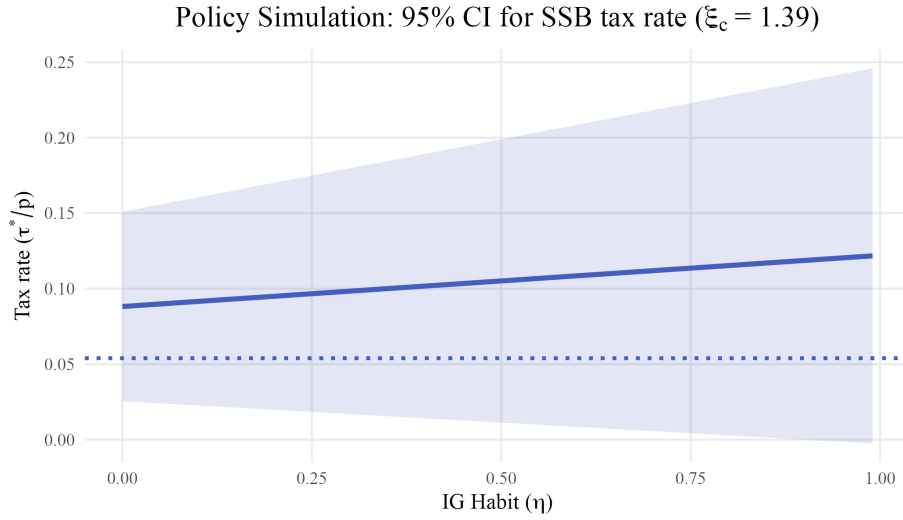
59

Confidence interval for Figure 3

Policy Simulation: 95% CI for SSB tax rate ($\xi_c = 1.39$)

Figure 5: With year and cohort FE 95% CI for the optimal SSB tax rate (τ^*/p) as a function of η ($\delta = 0.99$, $\xi_c = 1.39$)

Robustness Regarding FE selection

Without Cohort: column (7)

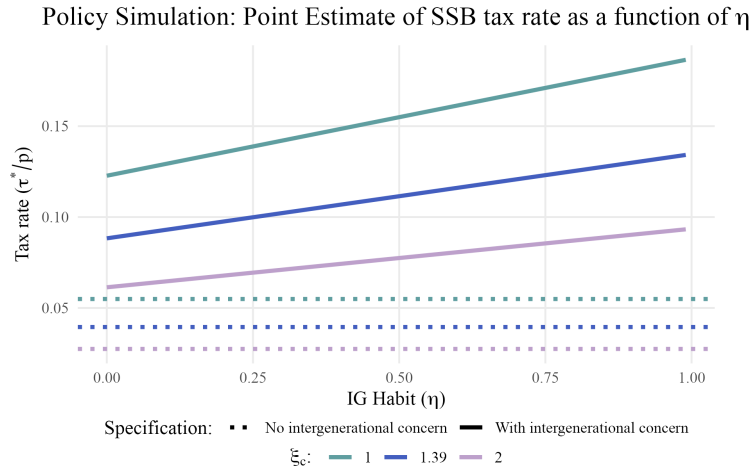Policy Simulation: Point Estimate of SSB tax rate as a function of η

Figure 6: Dropping Cohort FE Point estimate for the optimal SSB tax rate (τ^*/p) as a function of η ($\delta = 0.99$, $\xi_c \in \{1, 1.39, 2\}$)
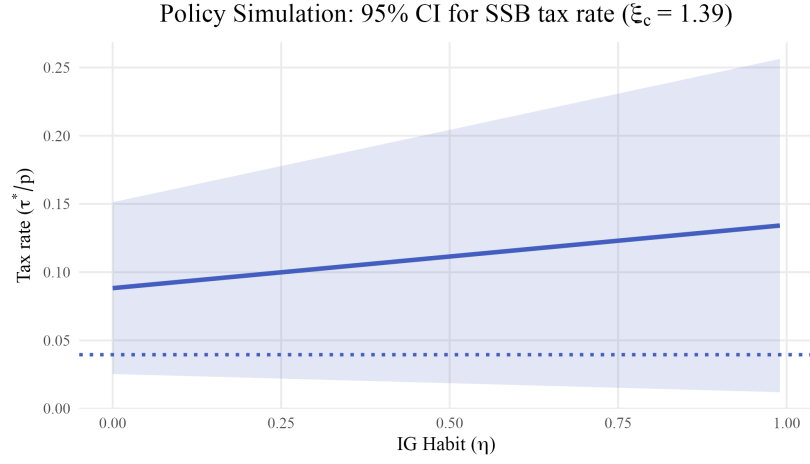
Figure 7: Dropping Cohort FE 95% CI for the optimal SSB tax rate (τ^*/p) as a function of η ($\delta = 0.99$, $\xi_c = 1.39$)

Elasticity	no IG	$\eta = 0$	$\eta = 0.25$	$\eta = 0.5$	$\eta = 0.75$	$\eta = 0.99$
1.00	0.055	0.123	0.139	0.155	0.171	0.187
1.39	0.040	0.088	0.100	0.111	0.123	0.134
2.00	0.027	0.061	0.069	0.077	0.086	0.093

Table 8: Dropping cohort FE Simulated optimal tax rate (τ^*/p) for selected elasticity and intergenerational habit term (η)
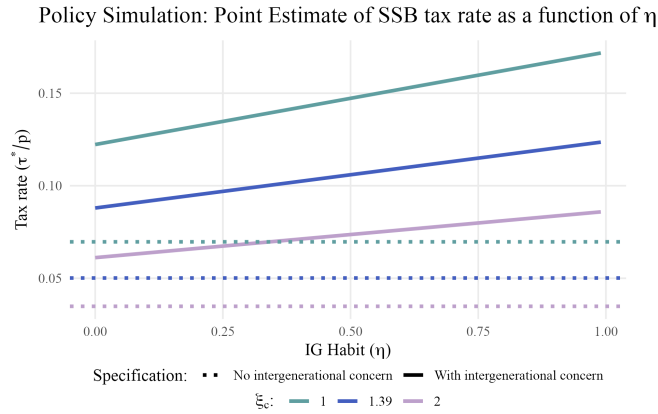
Without Year FE: column (6)

Figure 8: Dropping Year FE Point estimate for the optimal SSB tax rate (τ^*/p) as a function of η ($\delta = 0.99$, $\xi_c \in \{1, 1.39, 2\}$)
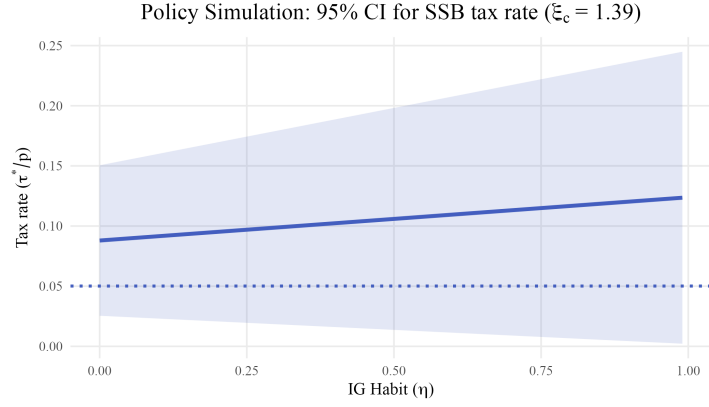
61

Policy Simulation: 95% CI for SSB tax rate (ξ_c = 1.39)

Figure 9: Dropping Year FE 95% CI for the optimal SSB tax rate (τ^*/p) as a function of η ($\delta = 0.99$, $\xi_c = 1.39$)

Elasticity	no IG	$\eta = 0$	$\eta = 0.25$	$\eta = 0.5$	$\eta = 0.75$	$\eta = 0.99$
1.00	0.070	0.122	0.135	0.147	0.160	0.172
1.39	0.050	0.088	0.097	0.106	0.115	0.124
2.00	0.035	0.061	0.067	0.074	0.080	0.086

Table 9: Dropping Year FE Simulated optimal tax rate (τ^*/p) for selected elasticity and intergenerational habit term (η)