CESifo WORKING PAPERS

6958
2018

March 2018

Too Slow a Change? Deep Habits, Consumption Shifts and Transitory Tax Policy

I.M. van den Bijgaart

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6958*
Category 1: Public Finance

Too Slow a Change? Deep Habits, Consumption Shifts and Transitory Tax Policy

Abstract

Habits are an important cause of sluggish consumption adjustment in response to price shocks. This paper studies shifts within the consumption bundle under endogenous habit formation. I put forward a model with good-specific, or 'deep', habits that cause persistence in good-specific consumption. In addition, at the aggregate level, habits act as a benchmark against which consumption is evaluated. I evaluate dynamic consumption choices under the realistic assumption that the consumer imperfectly internalizes the habit formation process. I compare consumption choices to the welfare-maximizing choices, and determine the path of taxes or subsidies that implements first-best consumption, both when goods are produced competitively and when they are produced by monopolists. I establish that a transition to a new consumption bundle is more likely inefficiently sluggish if the persistence effect is relatively strong. Strategic pricing behavior by monopolists leads to inefficiently rapid transitions. To explore the quantitative implications of the model I consider the introduction of a 10 percent charge on a subset of goods. I find that consumption adjusts inefficiently fast; implementing first-best adjustment requires a transitory discount of up to 60 percent of the cost increase.

JEL-Codes: D110, D620, H210, H230.

Keywords: habit formation, projection bias, consumption shifts, optimal taxation.

I.M. van den Bijgaart
University of Gothenburg
Vasagatan 1
Sweden – 40530 Gothenburg
inge.van.den.bijgaart@gu.se

March 12, 2018

1 Introduction

A rapidly expanding literature in behavioral economics documents how consumer preferences deviate from 'standard' neoclassical assumptions.[1] Preferences are shaped by context and reference points. In this paper, I consider the specific case of habit formation, where utility depends on habits in addition to consumption (Frederick and Loewenstein, 1999; Rabin, 2002). Here, habits shape demand, and as habits only slowly catch up with actual consumption, consumption patterns become persistent. People however have difficulty anticipating such changes in future preferences. This is known as projection bias, and implies individuals do not appropriately internalize the effects of current consumption decisions on future demand and welfare (Frederick and Loewenstein, 1999; Loewenstein et al., 2000; Conlin et al., 2007). Against this backdrop, a question arises whether, and in what way, consumption might deviate from the welfare-maximizing choice. If so, there might be room for fiscal policy to improve welfare by correcting the habit internality. This has relevance especially where policymakers anticipate or initiate a change in consumption patterns.

Over the past years, several major changes in consumption patterns have occurred, and continued change is anticipated in future decades. Many nations including Australia and California, have a long history of droughts, and as recently as 2014 suffered from acute water shortages. Yet, more intensive agriculture and population growth increase the difficulty of dealing with ensuing water shortages.[2] Combating water shortages and preventing an irreversible depletion of water resources required, and will continue to require, a substantial change in water consumption patterns by the agricultural sector and households.

Another illustration of a change in consumption patterns can be identified in past and future diets. Examples are numerous. Collapsing ocean fish stock will force consumers to shift their diets away from the most-prized species; so-called 'fat' and 'sugar' taxes have been proposed as a measure to induce consumers to adopt a healthier diet.[3] Similar examples can be found in other consumption categories. For instance, congestion and more stringent local pollution policies will require urbanites to abandon their gas-guzzling vehicle for a more efficient one, shift to public transport or even a bicycle. Stringent climate policies can contribute to this trend, by bringing an end to an era of cheap energy.

When habits cause consumption persistence, such shifts in consumption patterns will not fully materialise immediately. Rather, following a shock to prices, consumption, and habits, will only gradually adjust. In this context, the question is whether from a welfare perspective, such a shift in

[1] For example Rabin (2002), Tirole (2002), DellaVigna (2009) and Samson (2014).

[2] In California for instance, water shortages during the 2011-2017 drought led farmers to increasingly rely on the already dwindling groundwater stock, and reservoir water levels fell below 60 percent of average levels (State of California, 2015; The New York Times, 2015).

[3] See for instance The Washington Post (2012) and The Guardian (2016). Smed (2012) discusses the specific case of the Danish fat tax, which was introduced in 2012, yet later scrapped.

consumption patterns is too slow, or still too fast. Correspondingly, is a policy that further smooths this consumption adjustment welfare-improving, or should policy be used to implement a faster transition?

The examination of these questions requires a framework that captures how habits are commonly understood to shape demand and welfare. The examples above highlight the first effect; habits positively affect demand for specific goods. In addition, habits are viewed to affect welfare negatively; a higher habit 'stock' reduces utility from a given level of aggregate consumption.

I put forward a simple model which captures these effects of habit formation. In this model, a representative consumer forms habits at the level of individual goods. These good-specific habits cause persistence in consumption patterns. At the aggregate level, habits form a benchmark against which consumption is evaluated. This benchmark slowly adjusts to consumption, and causes the utility gain from increased consumption to fade over time; as the consumer gets used to a higher consumption level she loses (part of) her appreciation for it.[4]

This model setup offers two advantages over existing approaches. First, it reconciles two distinct approaches in the literature on habit formation; the aggregate habit approach (e.g. Abel, 1990; Monteiro et al., 2013) which considers only a single aggregate consumption good and the corresponding habit; and the 'deep' habits approach by Ravn et al. (2006; 2010), which allows for habit formation at the level of many specific goods. Second, existing specifications of good-specific habit formation (e.g. Ravn et al. 2006; 2010, Doi and Mino, 2008 and Nakamura and Steinsson, 2011) do not separate the multiple roles good-specific habits play; in these models the presence of good-specific habits affects (steady-state) welfare whenever habits lead to persistence in good-specific consumption, and vice versa. My specification is the first to explicitly separate these two roles, which allows for a closer evaluation of the importance of these roles and their relative strength in determining the optimal adjustment path of consumption.

I solve the model under the assumption that the consumer does not internalize that current consumption affects future habits and thereby future demand and welfare. Consumption decisions may therefore deviate from the optimal path, which is defined as the path that maximizes welfare, taking into account the endogenous formation of habits. I first consider the case where goods are produced competitively. I evaluate both the steady state, as well as the transition to a (new) steady-state consumption bundle, which might be initiated by a shock to the price of one good relative ot another.

I find that, as all goods are equally subject to habits, consumers optimally allocate their budget across goods in steady state. Hence, habits provide no reason to subsidize consumption of one good relative to another in steady state. Outside the steady state, habits do cause consumption decisions to deviate from their optimum. Whether the selected transition to a new consumption

[4] See for instance Carroll et al. (2000).

bundle is suboptimally slow or fast then depends on how important each role of habit formation is. Good-specific consumption persistence arises because the consumer prefers to consume goods she has a high habit in. In making her consumption decision however, the consumer does not internalize that current consumption affects future habits. As a consequence, she keeps 'too high' habits for those goods consumption shifts away from. A faster transition then improves welfare, as this will more rapidly bring down the habit of these goods. At the aggregate level the good-specific habits jointly act as a benchmark against which consumption is evaluated, with utility falling the higher this habit benchmark. Here, the transition offers an opportunity to temporarily reduce the habit benchmark. A slow transition, which implies the consumer consumes a relatively 'inefficient' bundle for a longer period of time, lowers this benchmark, and is thus beneficial. The optimal consumption path can be implemented by temporary, or transitory fiscal policy. A positive tax on those goods consumption shifts away from speeds up the transition, while a subsidy slows it down.

As market power is an important feature in many markets, I consider a second setting where goods are produced by monopolists instead. From the perspective of the forward-looking producers, habits constitute an investment in future sales. An anticipated drop in demand reduces the value of this investment, and increases the markup charged by monopolists. This price response speeds up the transition to the new consumption bundle compared to the competitive market. Hence, I find that while the laissez-faire transition might be inefficiently slow under perfect competition, it is always optimal to slow down transitions under monopolists. Hence, while taxes might still be called for under perfect competition, transitory subsidies are always required to implement the optimal consumption path under monopolists.

To illustrate the mechanisms and quantify effects, I evaluate the implications of an unanticipated shock to production costs. More specifically, I consider the introduction of a 10 percent charge on 'unhealthy foods', which induces consumption to shift away from these goods. Here I determine the transition when goods are produced by perfectly competitive firms or monopolists, as well as the optimal transition. I find the latter to be relatively slow; relative consumption drops by about 11 percent at the onset of the shift, and it takes more than 10 years for the economy to converge to the new steady state. Implementing this path requires sizable policy intervention; initial charges are set to about 40 to 60 percent of the long-run optimal charge. Appropriately managing this transition reduces transition costs by up to 5 percent. Immediately setting the charge at the long run level however has the advantage of being simple and straightforward to implement. I propose two alternative simple policy rules which generate welfare levels close to the one under the optimal path.

My paper contributes to the literature on habit formation and its implications,[5] and also the

[5] Early theoretical contributions on habit formation have been made by Pollak (1970) and Ryder and Heal (1973).

broader literature in 'behavioral public economics', which evaluates the policy implications of non-standard (behavioral) assumptions.[6] More specifically, I contribute to the literature that evaluates the implications of habit formation for optimal (tax) policy when consumers do not (fully) internalize the habit formation process. This imperfect internalization of the habit formation process opens room for welfare-improving policy intervention.[7] Ljungqvist and Uhlig (2000) for instance, show that in such cases, dynamic habit formation provides a rationale for procyclical taxes; such taxes counter the tendency to build up 'too high' habits during booms. Aronsson and Johansson-Stenman (2014) find that non-internalized habits tend to increase optimal marginal labor taxes, and have ambiguous effects on optimal capital taxes. In the context of growth, Alonso-Carrera et al. (2005), Turnovsky and Monteiro (2007) and Monteiro et al. (2013) characterize the income and consumption tax rates that implement the optimal path of consumption as the economy transitions to a balanced growth path. In Cremer et al.'s (2010) two-period model with retirement, habit formation and myopia cause overconsumption and undersaving in the first period of life. A tax on first-period consumption and a lump-sum transfer then implements the first-best allocation. If lump-sum transfers are unavailable, the second-best policy will also have redistributive implications. This theoretical research considers habits formed at the level of aggregate consumption instead of individual goods. Hence, this research cannot address the implications of habit formation for shifts *within* the consumption bundle. To my knowledge, I am the first to evaluate the potential policy implications of habit formation when habits are formed at the good-specific level. This is despite a large empirical literature commonly focusing on good-level habit formation.[8]

The distinction between aggregate and good-specific (deep) habits was first made by Ravn et al. (2006), who embed good-specific habits in a fully fledged dynamic general equilibrium model of the business cycle. Ravn et al. (2006), and later also Ravn et al. (2010) explore the implications of

---

The implications of habit formation have been explored in fields as diverse as asset pricing (Abel, 1990; Constantinides, 1990; Campbell and Cochrane, 1999), growth (Ryder and Heal, 1973; Carroll et al., 2000; Alvarez-Cuadrado et al., 2004; Alonso-Carrera et al., 2005), life cycle consumption and savings (Cremer et al., 2010; Koehne and Kuhn, 2014) and the relationship between income and happiness (Layard, 2006; Choudhary et al., 2012). Including habits in monetary policy and DSGE models has allowed these models to better capture certain features of the macroeconomy (Fuhrer, 2000; Ravn et al., 2006). Also empirical research generally confirms the presence of habit formation in consumption, see for instance Carrasco et al. (2005) and Bronnenberg et al. (2012).

[6] Examples of such behavioral assumptions include projection bias considered here, but also hyperbolic discounting, reference-dependent preferences, overconfidence, and limited attention (DellaVigna, 2009). See Bernheim and Rangel (2007) and Dalton and Ghosal (2011) for a general discussion.

[7] I consider the case of full projection bias. From a modeling perspective however, projection bias blurs the distinction between habit formation when habits are formed internally and own past consumption acts as a reference point, or externally, where the reference point depends on past consumption of a peer group (also known as 'catching up with the Joneses'). In both cases, external habits and internal habits with projection bias, the consumer does not internalize the habit formation process. For this reason, both the literature on, and policy implications of, internal and external habit formation are relevant to this paper and thus considered. This is further discussed in Section 2.

[8] For instance Carrasco et al. (2005); Zhen et al. (2011); Bronnenberg et al. (2012); Verhelst and Van den Poel (2014).

habit for monopolistic firms' pricing decisions in response to temporary shocks. One main result is that in the presence of market power, good-specific habits can give rise to excessive pass-through of cost shocks (Ravn et al., 2010). This paper builds upon their modeling approach, but examines a very different question, one that concerns the welfare implications of good-specific habits.

As discussed, consumption patterns can change for many reasons. Several of those reasons relate to resource scarcity and environmental externalities. In an extension, I show how my framework can be flexibly adjusted to include an environmental externality. In this extension, the temporary habit tax/subsidy can be reinterpreted as a time-varying environmental tax. Hence, this paper is the first to put forward habits as rationale for time-varying environmental taxes. This adds to the the more common rationales such as innovation externalities (Gerlagh et al., 2009; Acemoglu et al., 2012).

The remainder of this paper is structured as follows. The model is presented in Section 2. Section 3 discusses the equilibrium, including the steady state and the transition path towards this steady state. Optimal steady-state adjustment and policy are presented in Section 4, and Section 5 deals with the numerical application. Several of the model assumptions are discussed in Section 6; Section 7 concludes.

2 Model

In this section I put forward a structure that generalizes existing approaches to modeling habit formation. More specifically, it reconciles models of aggregate habit formation in the tradition of Abel (1990) and Carroll (2000), with the good-specific habits and consumption persistence as in Ravn et al. (2010). Habits play slightly different roles in these two frameworks. In models of aggregate habit formation, an increase in aggregate consumption C today, is linked to an increase in future habit stock H. These habits can be considered as a standard of living. Typically, the higher this standard of living H, the (i) lower utility for a given level of C, and (ii) higher marginal utility of consumption C , which in turn increases aggregate consumption demand. Mathematically, this requires $\frac{\partial U}{\partial H} < 0$ and $\frac{\partial^2 U}{\partial C \partial H} > 0$, respectively.[9] When considering habits at the level of specific goods, an additional role is added; (iii) the habit stock for some good i, h_i , the greater the demand for good-i consumption c_i (see for instance Ravn et al 2006; 2010). This again requires the marginal utility of consumption to be increasing in the habit, yet now at the good-specific level: $\frac{\partial^2 U}{\partial c_i \partial h_i} > 0$.

I consider a simple setup that satisfies these requirements, and also allows me to separate the good-level persistence effect, and aggregate 'standard-of-living' effect of habit formation. In this setup, a representative consumer consumes a variety of goods $i \in [0, 1]$. Time t good i consumption reads $c_i(t)$, and the consumer forms habits $h_i(t)$ over the same varieties. These good-specific, or

[9] Where I implicitly assume the utility function is increasing and concave in C.

'deep', habits cause persistence in consumption decisions: demand for good $c_i(t)$ is increasing in habit $h_i(t)$. Good-specific consumption and habits are aggregated into $C(t)$ and $H(t)$. The representative consumer's instantaneous utility $U(t)$ at time t increases in effective consumption $C(t)$, and $C(t)$ relative to a benchmark, the aggregate habit $H(t)$. The higher this benchmark, the lower utility from consumption. Hence, the aggregate habit causes some degree of hedonic adaptation: the utility gain from a permanent increase in consumption (partly) fades out over time as consumers become accustomed to the higher consumption level.[10]

Welfare reads

$$W(t) = \int_t^\infty e^{-\rho(v-t)} U(v) dv, \tag{1}$$

with instantaneous utility

$$U(t) = \frac{\left(C(t)^{1-\gamma}\left(\frac{C(t)}{H(t)}\right)^{\gamma}\right)^{1-\sigma}}{1-\sigma}. \tag{2}$$

The parameter γ is the aggregate habit strength, and measures the importance of the aggregate habit benchmark in utility. I allow for any $\gamma \in [0,1]$. Effective consumption is an aggregate of consumption over a variety of goods c_i. The importance of each variety in C depends on good-specific consumption weights w_i. These weights in turn depend on habits; a higher good-specific habit relative to the aggregate habit increases the weight of a good c_i in C:

$$C(t) = \left[\int_0^1 w_i(t) c_i(t)^{\frac{\eta-1}{\eta}} di\right]^{\frac{\eta}{\eta-1}}, \tag{3}$$

and

$$w_i(t) = \left(\frac{h_i(t)}{H(t)}\right)^{\frac{\theta}{\eta}}. \tag{4}$$

Here, η is the instantaneous elasticity of substitution across varieties and $\theta \in [0,1)$ is the good-specific habit strength. Deep habits, at the level of individual varieties, increase (relative) demand for specific varieties as they increase these varieties' weight in the consumption aggregate. Note that the aggregation from c_i to C preserves linear homogeneity: a proportional increase in all c_i translates into an equiproportional increase in C.

The aggregate habit is a measure for the effective consumption level the consumer is accus-

[10]The hedonic treadmill, or hedonic adaptation, is a concept from psychology which describes the tendency for humans to quickly return to a relatively stable level of happiness following a major positive or negative life event (Frederick and Loewenstein, 1999).

tomed to and defined as follows

$$H(t) = \left[\int_0^1 w_i(t) h_i(t)^{\frac{\eta-1}{\eta}} di \right]^{\frac{\eta}{\eta-1}}. \tag{5}$$

H is linearly homogeneous in good-specific habits h_i which implies consumption weights are independent of the scaling of the habit.[11] A proportional increase in all habits h_i thus affects utility only through an equiproportional increase in the aggregate habit benchmark H; it does not alter effective consumption C. Similarly, a shift in good-specific habits, keeping the aggregate habit H constant, affects utility only through its effect on the good-specific consumption weights w_i and effective consumption C. Such a shift in good-specific habits will increase effective consumption C if it brings the pattern of habits more in line with the pattern of consumption.[12]

Good-specific habits slowly catch up with consumption:

$$\dot{h}_i(t) = \xi \left(c_i(t) - h_i(t) \right), \tag{6}$$

where the dot denotes a time derivative and $\xi > 0$ is the adjustment speed of the habit. In steady state, habits have converged to actual consumption: $h_i = c_i$. From (3) and (5) it then also follows that in steady state, the aggregate habit equals effective consumption: $H = C$.

The above specification allows me to disentangle two roles of habits. First is the 'standard of living' role, captured by the γ. Consider a one-off equiproportional increase in all c_i (e.g. due to a permanent positive income shock). This increases C in (3) and hence utility in (2). As habits catch up, the H will increase over time, while the w_i will remain unchanged (due to the linear homogeneity of H in h_i). This increase in the aggregate habit H will in turn reduce the utility gain due to the consumption increase. If $\gamma = 1$, changes in consumption do not lead to long-term utility gains or losses (note $H = C$ in steady state), while for $\gamma = 0$, aggregate habits do not affect utility from C. In this case, (2) collapses to a standard CRRA utility function.[13]

The second role relates to persistence in good-specific consumption, which follows from a desire to bring the *pattern* of consumption in line with the *pattern* of habits. This role is captured

[11] I present (5) in current form to allow easy comparison to (3), even though through w_i, the variable $H(t)$ features in both the LHS and RHS of (5). It can however be easily verified that $H(t)$ is uniquely defined.

[12] This can be illustrated by the following example. Consider a consumption bundle that is high in vegetables and low in meat. Then effective consumption C derived from this bundle is higher if the consumer is used to this high vegetable, low meat diet, than if she were used to a low vegetable, high meat diet. Both diet habits however, could resemble the same standard of living, i.e. the same H.

[13] Jointly with σ, γ can also capture aggregate consumption persistence. As long as $\gamma > 0$ and $\sigma > 1$, $\frac{\partial^2 U}{\partial C \partial H} > 0$, i.e. the higher is H, the higher is the marginal utility from C and in turn demand for C. Whether this indeed leads to persistence in C then depends on whether C can be adjusted either through a labor-leisure, or consumption-savings decision. This paper abstracts from this adjustment channel. As discussed in Section 6, this abstraction is inconsequential for the analysis of within-bundle shifts, which is the focus of this paper.

by the consumption weights, with its strength determined by the θ. The higher is θ, the larger the consumption weight of a high-habit good relative to a low-habit good (see (4)), and hence the lower effective consumption C if the consumption pattern deviates from this pattern of habits (i.e. if c_i is high for i with low h_i and vice versa). As a consequence, the consumer will have a tendency to stick to previously-established consumption patterns, leading to consumption persistence. This effect is particularly strong for θ close to unity. If $\theta = 0$, all weights are equal to 1, and (3) collapses to a standard Dixit-Stiglitz specification with good-specific demand is independent of the good-specific habit.

For production, I assume a constant returns to scale production technology, where the production of each good requires $\delta_i > 0$ units of labor. In the introduction I provided several examples of changing consumption patterns. For some of consumption categories, such as water, marginal cost pricing is a reasonable assumption. For other categories however, for instance sugary drinks, goods are sold at substantial markups, indicating that market power plays a role. For this reason, throughout the analysis I consider both the case where goods are produced under perfect competition, as well as the setting where each good i is supplied by a monopolist.

I assume total labor supply, L, is fixed, so that the labor market equilibrium reads as follows:[14]

$$L = \int_0^1 \delta_i c_i(t) di. \tag{7}$$

In addition to the direct labor cost of production, producers may face a good-specific production tax. I denote the wage rate by p_L. To the producer, the total cost of producing one unit of c_i then equals $\delta_i p_L(t) \tau_i(t)$ where $\tau_i(t)$ is the gross tax rate; for $\tau_i(t) > 1$, good i production is subject to a positive tax, and good i is subsidized if $\tau_i(t) < 1$. Tax receipts are rebated though a lump-sum transfer (which is negative in case total receipts are negative).

Finally, in the remainder of the analysis I make two assumptions regarding the rationality of the consumer and producers:

Assumption 1 The representative consumer is subject to strong projection bias, i.e. it does not internalize the effect of current consumption on future habits.

Projection bias is a form of limited rationality where individuals do not (fully) anticipate future changes in preferences (Loewenstein et al., 2000; Samson, 2014). As a consequence, in the face of changing preferences, the individual is unable to fully optimize its consumption decisions. In the context of the current framework, this implies that the consumer does not internalize the habit

[14] I thus abstract from a labor-leisure tradeoff. This is to keep the analysis concise, and inconsequential for the paper's main results, which deal with the within-bundle consumption allocation. This is discussed in more detail in Section 6.

formation process (6). As a consequence, demand will be a function of the goods' current prices and habits, but not of future expected prices.

Assumption 2 Producers are forward-looking and atomistic.

Contrary to the consumer, producers do anticipate that current consumption affects future demand through habits. Hence, they adjust their optimization accordingly. Their atomistic size however implies that even though producers internalize the direct effect of $c_i(t)$ on the evolution of the good-specific habit $h_i(t)$, they do not internalize the subsequent effect on the aggregate habit $H(t)$.[15] As we will see below, habits affect the price charged by monopolists. If markets are perfectly competitive, goods are always sold at marginal costs.

Catching up with the Joneses In terms of modeling, the internal habit formation framework presented above is virtually equivalent to a setup which instead features 'catching up with the Joneses', as in Abel (1990), Alvarez-Cuadrado et al. (2004) and Alonso-Carrera et al. (2005). In such a setup, $h_i(t)$ represents an external habit, i.e., a reference point based on (past) consumption in the peer group. To the individual consumer, this habit is exogenous. With a representative consumer, one can show h_i still evolves according to (6).[16] In the remainder of the paper, I continue to interpret h_i as an internal habit where the consumer does not internalize the habit formation process. All results and policy recommendations continue to apply if habits are instead formed externally as described above.

3 Equilibrium

The representative consumer maximizes welfare (1) while taking habits as given. This gives demand

$$c_i(t) = \left(\frac{p_i(t)}{P(t)}\right)^{-\eta} \left(\frac{h_i(t)}{H(t)}\right)^{\theta} C(t), \tag{8}$$

[15]This implication is akin to the notion that monopolistically competitive firms internalize the effect of output on the good-specific price, but not on the aggregate price level in the economy.

[16]More specifically, let j be the indicator for the consumer, such that $c_{ji}(t)$ is the time t good i consumption of individual $j \in [0,J]$. Then (3) can be rephrased as $C_j(t) = \left[\int_0^1 w_i(t) c_{ji}(t)^{\frac{\eta-1}{\eta}} di\right]^{\frac{\eta}{\eta-1}}$. (4)-(6) still apply, where in (6) c_i is now redefined as $c_i(t) \equiv \int_0^J c_{ji}(t) dj$. In a representative consumer setting, $C_j(t)$ then collapses to $C(t)$ as in (3).

where p_i is the price of good i and

$$P(t) = \left[\int_0^1 \left(\frac{h_i(t)}{H(t)} \right)^{\theta} p_i(t)^{1-\eta} di \right]^{\frac{1}{1-\eta}} \tag{9}$$

is the price of effective consumption. Demand for good i decreases in the price of good i and increases in effective consumption C. For given aggregate habit H, a higher good-specific habit h_i increases the weight of consumption c_i in C (see (3)). As a consequence, demand increases in the good-specific habit. This in turn also increases the weight of the good i price in the price aggregate P. In the remainder, I take effective consumption as the numeraire, and thus normalize $P = 1$. Through the representative consumer budget constraint, expenditures C must equal income: $C(t) = \left[p_L(t)L + \int_0^1 \pi_i(t) di + \Omega(t) \right]$, where $\Omega(t) = \int_0^1 (\tau_i(t) - 1)\, \delta_i p_L(t) c_i(t) di$ is a lump-sum transfer.

Perfect competition On the producer side, price-setting is straightforward if markets are perfectly competitive. In this case we have

$$p_i(t) = \delta_i p_L(t) \tau_i(t). \tag{10}$$

Monopolists If instead the producers are monopolists in their respective goods, each chooses a series of prices that maximizes its firm value V_i, which is equal to the present value of profits, discounted at market interest rates: $V_i(t) = \int_t^{\infty} e^{-\int_t^{\nu} r(x)dx} c_i(\nu) \left[p_i(\nu) - \delta_i p_L(\nu) \tau_i(\nu) \right] d\nu$, with c_i given by (8). Producers anticipate that a reduction in the current prices not only increases current sales, but also, through habits, future demand and profits. Setting a low price to build habit can thus be considered an investment in future profits. Hence, habits are expected to reduce markups, which is confirmed by the following result for the monopolist pricing rule:[17]

$$p_i = \frac{\eta}{\eta - 1} \left[\delta_i p_L(t) \tau_i(t) - \xi \kappa_{h_i}(t) \right], \tag{11}$$

where

$$\kappa_{h_i}(t) = \int_t^{\infty} e^{-\int_t^{\nu} (r(x)+\xi)dx} \frac{\theta}{\eta} \frac{c_i(\nu)}{h_i(\nu)} p_i(\nu) d\nu, \tag{12}$$

and I require $\eta > 1$ to ensure positive steady-state markups. The standard monopolistic competition pricing rule now includes a habit discount, $\xi \kappa_{h_i}$. The size of this discount depends on the shadow value of the habit to the monopolist, κ_{h_i}, multiplied by the direct effect of an increase in consumption on the future habit, ξ. The monopolist sets a low price if investing in the habit is

[17] See Appendix A.1 for detailed derivations.

valuable, i.e. if the shadow value of the habit is high. This is the case if (future) demand is very sensitive to the habit (high $\theta c_i/h_i$) and prices are high (high p_i). A low elasticity of substitution η then implies markups are high, and a large share of this price constitute pure profits. Future returns are discounted at a rate $r+\xi$, where a higher discount rate reduces the shadow value of the habit. This is due to the fact that a low persistence of the habit (high ξ) reduces the marginal effect of an increase in c_i today on habits further in the future, while higher interest rates reduce the present value of a given flow of returns.

Under monopolists, habits not only lead to lower markups, but also to time-varying markups. This can be seen as follows. Suppose that p_i is constant, and we initially have $h_i < c_i$. Then as habits catch up with consumption, c_i/h_i falls and so does the shadow value of the habit. This increases the monopolist's price according to (11) and is thus inconsistent with the constant price just assumed. The interest rate r is determined endogenously by the consumption Euler equation, which ensures optimal smoothing of (expected) marginal utility over time: $r(t) = \rho - \mathbb{E}^{rc}\left[\frac{\partial}{\partial t}\left(\frac{\partial U(t)}{\partial C(t)}\right) / \frac{\partial U(t)}{\partial C(t)}\right]$, where $\mathbb{E}^{rc}[\cdot]$ is the expectation from the perspective of the representative consumer.[18]

3.1 Steady state

The economy is in steady state if prices, consumption and habits are constant over time. Then, by (6), for all goods $i \in [0,1]$, habits must equal consumption: $c_i^* = h_i^*$. Here, the star indicates we are in steady state. In addition, by (3) and (5), it follows that in steady state also the aggregate habit equals effective consumption: $C^* = H^*$, and the market interest rate equals the rate of time preference ρ.

Perfect competition The good i steady-state price under perfect competition equals

$$p_i^* = \delta_i p_L^* \tau_i^*. \tag{13}$$

Monopolists To the monopolist, the steady-state shadow value of the habit is

$$\kappa_{h_i}^* = \frac{1}{\rho+\xi}\frac{\theta}{\eta}p_i^*, \tag{14}$$

[18]Since none of the theoretical results in the remainder of the text rely on whether the consumer's expectations regarding the evolution of H are rational or not, I make no further assumptions here. If the consumer does not anticipate the evolution of H, which is the assumption I consider most in line with Assumption 1, then $r(t) - \rho = \sigma\dot{C}(t)/C(t)$.

which with (11) gives the following steady state price:

$$p_i^* = \delta_i p_L^* \tau_i^* \frac{\eta}{\eta - 1} \left[\frac{\rho + \xi}{\rho + \xi \left(1 + \frac{\theta}{\eta - 1}\right)} \right]. \tag{15}$$

Even though habits reduce the monopoly markup in steady state, the markup remains positive.[19]

As the total labor supply is exogenous, consumption decisions are fully determined by *relative* prices. In the remainder of the paper, for ease of exposition and to stress this point, I will mostly focus on prices and quantities of a good i relative to some 'base' good $b \in [0,1]$. From (8), steady-state relative consumption then reads

$$c_i^{R*} = \left(p_i^{R*}\right)^{-\frac{\eta}{1-\theta}}, \tag{16}$$

with $c_i^R \equiv c_i/c_b$ and $p_i^R \equiv p_i/p_b$. The steady-state relative price is independent of market structure:

$$p_i^{R*} = \delta_i^R \tau_i^{R*}, \tag{17}$$

where δ_i^R and τ_i^R are defined in line with c_i^R and p_i^R. The steady state is interior and unique only if demand is strictly concave in the good-specific habit, i.e. only if $\theta < 1$. This condition is easily derived from (8). For a given set of prices, consumption scales with the habit at degree θ. If relative consumption, c_i^R, rises by 1 percent, future habits follow, and in turn future relative consumption goes up by an additional θ percent. The long run increase in c_i^R is then bounded only if $\theta < 1$.

This observation is mirrored in the result for the long run price elasticity of demand. With good-specific habit formation, the long run price elasticity of demand exceeds the short run one. This can be seen by comparing equations (8) and (16). From (8), the (absolute value of) the short run price elasticity of demand is equal to the instantaneous elasticity of substitution across goods: $\varepsilon_p^{SR} = \eta$. In the long run, this price elasticity of (relative) demand is $\varepsilon_p^{LR} = \eta/(1-\theta)$ (see (16)). In the absence of good-specific persistence ($\theta = 0$) these elasticities are equal. For positive θ, the long run shift in consumption in response to a permanent change in relative prices exceeds the short run one: $\varepsilon_p^{LR} > \varepsilon_p^{SR}$. If $\theta = 1$, ε_p^{LR} is unbounded, implying that, in the long run, goods act as perfect substitutes. As a consequence, not all goods may be consumed in steady state, and the relevant steady state would depend on initial values of the h_i. As stated in Section 2, I assume $\theta \in [0,1)$, which rules out such indeterminacy.

Finally, in steady state, due to uniform markups, the relative price p_i^R is independent of whether goods are produced under perfect competition or by monopolists. Outside of steady state however,

[19]The monopoly markup is positive if $\frac{\eta}{\eta-1}\frac{\rho+\xi}{\rho+\xi\left(1+\frac{\theta}{\eta-1}\right)} > 1$. This condition can be rearranged to $\rho > \xi(\theta - 1)$. As $\theta < 1$ and ρ, $\xi > 0$, this condition is always satisfied.

the relative price set by monopolists diverges from the perfect competition price ratio (see (10) and (11)). As will be shown in Section 3.2.2, this gives two distinct transition paths of consumption towards a steady-state equilibrium.

3.2 Transition

Due to habit formation, consumption might diverge from its steady state. When consumption lies above or below the habit, the habit will change over time, affecting future demand and possibly prices. Starting in a steady state, any shock to production costs, either through a shock to the unit labor requirement δ_i, or a permanent change in taxes τ_i, will cause consumption to deviate from the habit.

Consider for instance a permanent increase in the cost of energy. This cost shock could be due to the introduction of an economy-wide carbon tax, a shutdown of coal or nuclear power plants, or import restrictions on oil or gas implemented for geopolitical reasons. Also developments unrelated to a particular country's policies, such as increased global energy demand, or the depletion of oil and gas reserves will likely confront consumers with higher prices for energy. Such a permanent increase in energy cost, and correspondingly steady-state prices for energy-intensive goods, induces consumption to shift away from these goods. With slow habit adjustment, consumption of energy-intensive goods will then fall short of the habit and the consumption may require time to fully adjust to the lower-energy bundle. A qualitatively similar pattern will be observed in other applications. For instance, the introduction of a 'fat tax' will induce a substitution away from fatty foods.

When goods are produced under perfect competition, equilibrium prices are straightforwardly determined also outside the steady state. From (11), this is however not the case in the presence of monopolists; prices are a complex function of future consumption, habits and prices. To analytically evaluate and compare the transition paths of consumption, prices and habits towards their respective steady states I perform a first-order approximation of these paths in two steps. First, Section 3.2.1 provide a general characterization of the paths of consumption, prices and habits as the economy converges to the steady state. Then in Section 3.2.2, I use this characterization to evaluate changes in consumption in response to a permanent change in unit production costs. This characterization is also later used in Section 4.2, to evaluate the optimal path, and solve for the policy required to implement it.

3.2.1 General characterization

To approximate the path of consumption and prices I loglinearize the system around its steady state. Let a tilde denote a log-deviation from the steady-state, such that $\tilde{z}(t) \equiv dz(t)/z^* \approx (z(t) - z^*)/z^*$

and thus $\tilde{z}_i^R = \tilde{z}_i - \tilde{z}_b$ for some variable z. The loglinearized the demand equation (8) then reads

$$\tilde{c}_i^R(t) = -\eta \tilde{p}_i^R(t) + \theta \tilde{h}_i^R(t). \tag{18}$$

From (6), $\tilde{h}_i^R(t)$ evolves according to

$$\dot{\tilde{h}}_i^R(t) = -\xi \lambda \tilde{h}_i^R(t), \tag{19}$$

where I define the following linear relationship between $\tilde{c}_i^R(t)$ and $\tilde{h}_i^R(t)$:[20]

$$\lambda \equiv 1 - \tilde{c}_i^R(t)/\tilde{h}_i^R(t). \tag{20}$$

Then (18)-(20) give the following solutions for the evolution of relative consumption, prices and habits:

$$\tilde{c}_i^R(t) = [1-\lambda]\,\tilde{h}_i^R(t); \tag{21}$$

$$\tilde{p}^R(t) = \left[\frac{\theta - 1 + \lambda}{\eta}\right]\tilde{h}_i^R(t); \tag{22}$$

$$\tilde{h}_i^R(t) = \tilde{h}_i^R(0)e^{-\xi\lambda t}. \tag{23}$$

The variable $\tilde{h}_i^R(0)$ represents the initial deviation of relative habits from the steady state. Whenever this ratio of good i to b habits lies above its steady-state ratio $\tilde{h}_i^R(0) > 0$, while $\tilde{h}_i^R(0) < 0$ if the opposite applies.[21] Then for a given value of $\tilde{h}_i^R(0)$, the paths of consumption and prices are fully determined by the familiar parameters θ, η and ξ, and λ, the convergence factor. This convergence factor can be interpreted in two ways. First, λ, multiplied by the habit adjustment speed ξ, is the rate at which habits converge to the new steady state. The larger λ, the more rapid convergence. Second, λ determines the choice of consumption c_i for a given level of our state variable, the (relative) habit. The larger the convergence factor λ, the closer good i consumption will be to its steady state for a given steady-state deviation of habits. Of course, the two interpretations are interrelated. Current consumption affects future habits, which in turn adjust more rapidly the

[20] For ease of exposition, I implictly assume λ is constant and strictly positive. In Sections 3.2.2 and 4.2 I use loglinearized pricing rules to determine λ and find that λ is indeed constant and positive.

[21] Note that $\tilde{h}_i^R(0) = 0$ does not neccesarily imply all h_i are in steady state. For instance, suppose that we start in a steady state, and all goods i are hit by the same proportional shock to δ_i. This affects the steady-state levels of the c_i. Steady-state *relative* consumption and habits however are unaffected (see (16) and (17)). Hence, $\tilde{h}^R(0) = 0$ and consumption c_i will immediately jump to the new steady state, while habits h_i slowly adjust.

further is consumption from the habit. Hence, one should expect convergence to be fast if c_i^R is close to the steady state for a given h_i^R. Both the former, fast convergence, and the latter, $\tilde{c}_i^R$ close to zero, are indeed the case if λ is high. In the next two sections, I solve for the convergence factor under perfect competition and monopolistic supply respectively.[22]

3.2.2 Equilibrium transition

To determine the equilibrium transition, I solve for this convergence factor λ under perfect competition and monopolistic supply. For now, I assume taxes are constant, i.e. no additional intervention will take place along the transition. Section 4.2 assesses the first-best transition path, and determines the good-specific taxes and that implement this path. For ease of exposition I explain results in the context of a sudden and permanent increase in relative unit production cost $\delta_i^R \tau_i^R$. Starting in a steady-state, such a shock triggers a transition of consumption from good i to the base good b. In line with the examples discussed before, this good i may represent an energy-intensive good or unhealthy food.

Perfect competition Under perfect competition, prices adjust one-for-one with marginal costs (see (10)), which gives

$$\tilde{p}_i^R(t) = \tilde{\tau}_i^R(t). \tag{24}$$

With constant taxes $\tilde{\tau}_i^R = \dot{\tilde{\tau}}_i^R = 0$. This implies $\tilde{p}_i^R = 0$ at all times. From (22) I can thus determine the value for the convergence term λ^{pc}:

Lemma 1. $\lambda^{pc} = 1 - \theta$

Proof. In text. □

Then, from (21) and (23), I can conclude the following regarding the path of (relative) consumption and habits. An increase in relative marginal cost for good i causes good i consumption to fall relative to good b. In response to the drop in c_i^R, the relative habit falls. This induces a further decrease in relative consumption until the economy has converged to its new steady state. The transition to the steady state will be faster the faster is habit adjustment, i.e. higher ξ. Also, a low habit persistence (low θ) implies that both the initial drop in consumption is larger, and the economy transitions more rapidly to the new steady state.

[22] The paper focuses on gradual transitions, e.g. in response to a shock to production costs. Equation (21) can however also be interpreted in the context of a consumption or production quota. If the quota is binding, consumption immediately jumps to the steady state: $\tilde{c}_i^R(t) = 0$ for all t. This gives $\lambda = \lambda^{qt} = 1$.

Monopolists Under monopolistic supply, the current price is a complex function of future consumption, prices and habits. The linearization of the monopolist's pricing rule as expressed by (11) and (12) around the post-tax steady state gives[23]

$$\tilde{p}_i^R(t) = \tilde{\tau}_i^R(t) - \frac{\xi\theta}{\rho(\eta-1)+\xi(\eta-1+\theta)}\left[\tilde{c}_i^R(t) - \tilde{h}_i^R(t)\right] + \frac{\eta-1}{\rho(\eta-1)+\xi(\eta-1+\theta)}\dot{\tilde{p}}_i^R(t) - \frac{1}{\rho+\xi}\dot{\tilde{\tau}}_i^R(t). \tag{25}$$

Then for $\tilde{\tau}_i^R(t) = \dot{\tilde{\tau}}_i^R(t) = 0$, the following can be established regarding λ^{mc}:[24]

Lemma 2. *If $\theta > 0$, then $\lambda^{mc} \in \left(\lambda^{pc}, 1+\theta(\eta-1)^{-1}\right)$. If $\theta = 0$, then $\lambda^{mc} = 1$.*

Proof. See Appendix A.2.1 □

Lemma 3. *If $\theta > 0$, then $\frac{\partial\lambda^{mc}}{\partial\rho} < 0$, $\frac{\partial\lambda^{mc}}{\partial\xi} > 0$, and $\frac{\partial\lambda^{mc}}{\partial\eta} < 0$. For any θ, $\frac{\partial\lambda^{mc}}{\partial\theta} > 0$ iff $\eta - 1 < (1-2\theta)\xi/(\rho+\xi)$. Finally, $\lambda^{mc} > 1$ iff $\theta > 0$ and $\eta - 1 < (1-\theta)\xi/(\rho+\xi)$.*

Proof. See Appendix A.2.2 □

Whenever habits cause persistence at the level of specific goods ($\theta > 0$) convergence is faster if goods are produced by monopolists instead of perfectly competitive firms. This is due to the fact that it is optimal for producers to increase relative prices in *excess* of the marginal cost increase. When the positive marginal cost shock hits, the producer of a *i*-good realizes that future demand for *i* falls below the current habit. This reduces the return to investment in the habit. In response, the producer increases the markup, and thus increase prices by more than the increase in unit production costs. For a good-*b* producer, the exact opposite story holds: it anticipates an increase in (relative) demand, which increases the return to investment in the habit. A good-*b* producer thus chooses a lower markup than its good-*i* competitor. As consumption converges to the new steady state, the relative price will fall toward the long run relative price, which is equal to the ratio of marginal costs.

As initially, relative prices increase by more than the increase in relative marginal costs, the drop in relative consumption under monopolistic supply is greater than the one under perfect competition. In fact, the shift in prices may be so large that c_i^R undershoots the long run equilibrium. One can show this is the case if $\theta > 0$ and $\eta - 1 < (1-\theta)\xi/(\rho+\xi)$. This latter condition is more likely satisfied if goods are weak substitutes (η is low), habits are weak yet change rapidly (low θ and high ξ) and time preference is weak (low ρ). The intuition is subtle, and relates to the sensitivity of prices to good-specific habits, compared to the sensitivity of consumption to these

[23] Detailed derivations can again be found in Appendix A.1.

[24] The closed-form solution for λ^{mc} can be found in Appendix A.3.1.

habits, taking prices as given. Suppose that the good i habit is above its steady state ($\tilde{h}_i^R(0) > 0$). Then from (12), this causes a large drop in the shadow value of the habit, κ_{h_i}, if habits affect future demand rapidly (high ξ), future returns are discounted little (low ρ), the elasticity of demand, η, is low and demand is sensitive to the habit (high θ). This drop in κ_{h_i} increases the monopolist's price p_i, which reduces c_i. The high θ however also implies consumption responds strongly to the above steady-state habit. This outweighs the effect of θ through prices; with a high θ, c_i^R is less likely to undershoot the long run equilibrium if $\tilde{h}_i^R(0) > 0$. More generally, and consistent with the intuition above, whenever $\theta > 0$, λ^{mc} is increasing in ξ, and decreasing in ρ and η. The effect of a change in θ on λ^{mc} is ambiguous and depends on parameter values.[25]

4 Optimal consumption and implementation

A change in relative prices always induces the consumer to reconsider its consumption choices and, over time, shift to a new consumption bundle. The consumer however is not perfectly rational. She is subject to projection bias and thereby does not internalize the effect of current consumption on future preferences through habit formation. Consumption choices are thus likely suboptimal; from a welfare perspective, the consumer may adjust her consumption choices too slowly, or too rapidly. In this section, I determine optimal consumption and prices, both in the steady state and along the transition towards the steady state. Optimal consumption choices are defined as the paths of consumption, $[c_i(\nu)]_{i\in[0,1],\nu=t}^{\nu=\infty}$, that maximize the present value of instantaneous utilities (2), subject to (3)-(5), taking into account the endogenous formation of habits (6), and labor market equilibrium (7). I then assess how policy can be used to implement this first-best consumption allocation. More specifically, I solve for the (path of) good i taxes or subsidies that induce optimal consumption choices.[26] As established in the previous section, the transition path without intervention depends on the underlying market structure. In line with this result, the tax path that implements optimal consumption choices under perfect competition differs from the one when goods are produced by monopolists.

[25] Lemmas 2 and 3 can be considered a generalization of a result presented in Ravn et al. (2006), and discussed in more detail by Ravn et al. (2010). This result states that monopolistic producers may increase markups following a temporary positive marginal cost shock. An increase is more likely the more persistent the shock, and for the limiting case where the shock is fully persistent, producers always increase markups. Ravn et al. (2010) arrive at this result in a discrete-time framework where $h_{it} = c_{it-1}$. Lemma 2 generalizes this result to a continuous time setup with slow habit adjustment and a permanent shock. Lemma 3 then points at the novel result that consumption may undershoot its long-run equilibrium.

[26] I focus on the use of taxes and subsidies to implement the first-best allocation. As the model features no uncertainty, any allocation implemented by a given path of taxes/subsidies can also be implemented by (time-varying) quota. Referring to Dalton and Ghosal (2011), this implies I take an (in)direct paternalistic approach to policy intervention where I implicitly assume the policymaker has full information regarding preferences and their evolution over time. I thus do not consider a soft-libertarian approach, where policy would take the form of teaching the consumer to internalize the endogenous habit formation process herself.

As the previous section, this section abstracts from the source of the increase in relative production cost $\delta_i^R \tau_i^R$, which sets in motion a transition to a new steady state. In Appendix B I extend the model to account for the presence of a (positive or negative) externality due to the production or consumption of one or multiple goods (e.g. an environmental externality). In this case the correction of such an externality affects the production costs of a subset of firms. I show that all results concerning the model dynamics carry through.[27]

Welfare (1) is maximized subject to (2)-(7). I solve the Hamiltonian and use consumer demand (8) to arrive at the following rule for optimal prices:[28]

$$p_i(t) = \delta_i \check{\mu}_L(t) - \xi \check{\mu}_{h_i}(t), \tag{26}$$

with

$$\check{\mu}_{h_i}(t) = \int_t^\infty e^{-(\check{r}+\xi)(\nu-t)} \frac{c_i(\nu)}{h_i(\nu)} p_i(\nu) \left[\frac{\theta}{\eta-1} - \left[\gamma + \frac{\theta}{\eta-1}\right] \left(\frac{h_i(\nu)}{H(\nu)}\right)^{\frac{\eta-1}{\eta}} \left(\frac{c_i(\nu)}{C(\nu)}\right)^{-\frac{\eta-1}{\eta}} \right] d\nu, \tag{27}$$

where μ_L is the shadow value of labor and μ_{h_i} the shadow value of the habit to the consumer in terms of C, and $\check{r} = \rho - \left[\frac{\partial}{\partial t}\left(\frac{\partial U(t)}{\partial C(t)}\right) / \frac{\partial U(t)}{\partial C(t)}\right]$. The optimal price for c_i equals its marginal production cost, minus the marginal value of c_i due to habit formation. This value is equal to the direct effect of an increase in c_i on the future habit, ξ, multiplied by the shadow value of the habit, $\check{\mu}_{h_i}$. The shadow value of the habit captures the effect of an increase in h_i on future welfare and can be separated into two components. First, for $\theta > 0$, an increase in the good-specific habit increases the consumption weight w_i, which increases the benefit from c_i in C(see (3)). Simultaneously however, through H, an increase in h_i reduces the weight of all other goods. The net effect on C is positive only if c_i/C is large compared to h_i/H. Put differently, an increase in h_i has positive value if it brings the 'pattern' of habits (h_i/H) more in line with the 'pattern' of consumption (c_i/C).

The second component is captures the welfare effect of the aggregate habit benchmark and is negative whenever $\gamma > 0$. Any increase in the good-specific habit h_i increases the aggregate habit H. This rise in the consumption benchmark in turn reduces utility for a given level of effective consumption C.

[27]More specifically, any optimal policy intervention can be decomposed into two parts: i) state-independent (Pigovian) taxes that correct for the production externality. ii) state-dependent taxes that manage the rate at which consumption substitutes away from c_i, as discussed in the main part of this paper. See Appendix B for more details.

[28]See Appendix A.1 for detailed derivations.

4.1 Steady state

In steady state, consumption equals habits, both at the good-specific and the aggregate level. This in turn implies prices are constant. From (27), the steady-state shadow value of the habit is

$$\check{\mu}_{h_i}^* = -\gamma \frac{1}{\rho + \xi} p_i^*, \tag{28}$$

which, with (26), gives the following solution for the optimal steady state good i price:

$$p_i^* = \delta_i \check{\mu}_L^* \left[\frac{\rho + \xi}{\rho + \xi (1 - \gamma)} \right]. \tag{29}$$

Whenever $\gamma > 0$, the shadow value of the habit is negative in steady state. Whereas good-specific persistence is not associated with any steady-state welfare effects, the aggregate habit causes a negative long-run effect on utility which the consumer does not internalize. The larger γ, the greater this negative externality on the future self (i.e. negative internality), which translates into a higher steady-state markup. More rapid adjustment of consumption to the habit implies the externality occurs sooner. Like a lower time preference, this increases the present value of the internality and thereby the optimal steady-state markup.

I can then establish the following:

Proposition 1. *In steady state, laissez-faire consumption choices are optimal. Any uniform tax implements this optimum.*

Proof. By (29), the optimal relative price in steady state satisfies $p_i^{R*} = \delta_i^R$. This is equal to (17) with $\tau_i^{R*} = 1$. Under laissez faire, $\tau_i = 1$ for all i, so $\tau_i^{R*} = 1$. □

As habits and market power affect demand and supply of all goods to an equal extent, they do not distort the steady-state allocation of consumption across goods. Hence, habits do not provide a rationale for taxing or subsidizing one good more aggressively than another in the long run. As pointed out before, in the absence of savings and with inelastic labor supply, consumption decisions are fully determined by relative prices. As a consequence, any tax that is uniform across goods, including zero taxes, implements this first-best allocation.[29]

[29] If we would extend the model to include endogenous labor supply such as in Cremer et al. (2010), or allow the consumer to transfer consumption across time, as in Carroll et al. (2000), price and tax level changes would affect consumption levels. Now, due to noninternalized habits, the steady-state consumption level is likely inefficient. In such a case, from (13), (15) and (29), a steady-state habit tax equal to $T^* = \frac{\rho+\xi}{\rho+\xi(1-\gamma)}$ and $T^* = \frac{1}{\eta} \frac{(\rho+\xi)(\eta-1)+\xi\theta}{\rho+\xi(1-\gamma)}$ under perfect competition and monopolistic supply respectively implements the first-best steady-state consumption (note I implicitly assume the equilibrium wage is equal to $\check{\mu}_L$). Note that results for consumption, price and tax *ratios* are independent of the *levels* of these variables, and thus independent of assumptions regarding labor supply and savings. See also the discussion in Section 6.

As will be demonstrated in the next subsection, this result only holds in the steady state. Along the transition towards the steady state, taxes and subsidies may be required to implement optimal consumption choices.

4.2 Transition

To determine the optimal path of consumption, prices and habits , I adopt the same approach as in Section 3.2, where I solved for the convergence factor λ under perfect competition and monopolistic supply. With (21) and (22), this convergence factor pins down the paths of consumption and prices outside the steady state. To find the λ for the optimal path, λ^{opt}, I first linearize (26) and (27) to find

$$\tilde{p}_i^R = -\frac{1}{\eta}\frac{\xi(\theta-\gamma)}{\rho+\xi(1-\gamma)}\left[\tilde{c}_i^R-\tilde{h}_i^R\right]+\frac{1}{\rho+\xi(1-\gamma)}\dot{\tilde{p}}_i^R. \tag{30}$$

I can then establish the following regarding λ^{opt}:[30]

Lemma 4. *If $\gamma \neq \theta$, then $\lambda^{opt} \in (\min\{1-\gamma, \lambda^{pc}\}, \max\{1-\gamma, \lambda^{pc}\})$ while if $\gamma = \theta$, then $\lambda^{opt} = \lambda^{pc} = 1-\theta$. Next, $\lambda^{opt} < \lambda^{mc}$ if $\max\{\gamma, \theta\} > 0$, while $\lambda^{opt} = \lambda^{mc} = 1$ if $\gamma = \theta = 0$.*

Proof. See Appendix A.2.3 □

Lemma 5. *For any γ and θ, $\frac{\partial \lambda^{opt}}{\partial \gamma} < 0$, $\frac{\partial \lambda^{opt}}{\partial \theta} < 0$, and $\frac{\partial \lambda^{opt}}{\partial \eta} = 0$. $\frac{\partial \lambda^{opt}}{\partial \rho} > 0$ and $\frac{\partial \lambda^{opt}}{\partial \xi} < 0$ if $\gamma > \theta$, while $\frac{\partial \lambda^{opt}}{\partial \rho} < 0$ and $\frac{\partial \lambda^{opt}}{\partial \xi} > 0$ in case $\gamma < \theta$, and $\frac{\partial \lambda^{opt}}{\partial \rho} = \frac{\partial \lambda^{opt}}{\partial \xi} = 0$ if $\gamma = \theta$.*

Proof. See Appendix A.2.4 □

From which follows

Proposition 2. *Suppose goods are produced under perfect competition. If $\gamma > (<)\theta$, the laissez-faire transition to the steady state is suboptimally fast (slow). The optimal adjustment path can then be implemented by introducing a positive and declining subsidy (tax) on good i whenever $\tilde{h}_i^R(0) > 0$ and a positive and declining tax (subsidy) on good i if $\tilde{h}_i^R(0) < 0$. If $\gamma = \theta$, the laissez-faire transition to the steady state is optimal, and any constant subsidy (tax) implements this path.*

Proof. See Appendix A.2.5 □

Suppose we are again in the situation where e.g. due to a price shock $\tilde{h}_i^R(0) > 0$; the good *i* relative habit lies above its steady state. Then if we take the transition where consumers face a flat price schedule with $\tilde{p}_i^R = p_i^{R*}$ as a benchmark, it is optimal to speed up the transition from good *i* to

[30] The closed-form solution for λ^{opt} can be found in Appendix A.3.1.

b if the good-specific habit parameter θ is larger than the aggregate habit parameter γ, whereas the opposite holds if $\gamma > \theta$. This result can be explained as follows. The consumer does not internalize the effect of current consumption on future habits. These habits however do affect future utility through the consumption weights w_i and the aggregate habit H. Whether a slower or faster shift in consumption from i to b is welfare-improving then depends on whether a slower or faster shift in habits increases future utility through w_i and H.

Starting with the effect through w_i, I find that a faster transition is welfare-improving. This can be seen as follows. An increase in $\tilde{p}_i^R$ induces the consumer to shift consumption away from good i and towards good b. This shift causes a larger increase (smaller drop) in future effective consumption C the higher is the weight of good b relative to good i. Hence, future effective consumption C increases if the weight of good b, relative to good i, rises. This can be achieved by building habit in good b, and divesting habit in i, which is in turn requires consumption to more rapidly shift away from good i and towards good b. To summarize, building b habit is beneficial if b consumption is rising, and conversely, a relatively high i habit is costly if good i consumption is falling. Hence, welfare is increased if consumers more rapidly get rid of this i habit.

Second, good-specific habits negatively affect welfare as through H, they jointly act as a benchmark against which effective consumption is evaluated. The transition offers an opportunity to manage, i.e. temporarily reduce, this benchmark H. It turns out this argues in favor of a slow transition away from good i consumption. Although not immediate, the result is intuitive. At each point in time, the consumer chooses i and b consumption such that it maximizes effective consumption C. Following an increase in the relative price for good i, consumption shifts away from this good, as postponing, or slowing down this shift, would reduce effective consumption C. A slow transition however also has an advantage, as 'too high' consumption of the now relatively expensive good pulls down the reference habit H.[31]

If $\theta = 0$, the consumption weights w_i are independent of habits and hence habits do not cause good-specific consumption persistence. This implies the first effect is absent, and only the second effect, arguing in favor of a slower transition, is relevant. Similarly, if $\gamma = 0$, the benchmark H does not affect utility for given C, and habits only affect future utility through w_i. More generally, which of the two mechanisms dominates depends on whether habits are stronger at the good-specific or at the aggregate level. This can be evaluated by a simple condition comparing the deep habit strength, θ, to the aggregate habit strength, γ, as described in Proposition 2.

[31] As an extreme example, think of the following. Suppose consumption consists of apples and oranges. Then a strong increase in the price of apples initiates a shift towards oranges in the consumption bundle. Suppose the consumer is stubborn, and initially sticks to an apple-intensive diet. Since apples are very expensive, the consumer can afford only a few and is very hungry. The next period, the consumer decides to spend less on apples such that he can buy many oranges. As the consumer was used to starving in the previous period (H dropped a lot), the increase in orange consumption and elimination of hunger constitutes a large welfare gain.

Finally λ^{opt}, and hence the optimal adjustment rate, falls in θ and γ. The effect of a change in ρ or ξ depends on whether λ^{opt} is larger or smaller than λ^{pc}. If $\lambda^{opt} > \lambda^{pc}$, λ^{opt} is decreasing in ρ and increasing in ξ.[32] If $\lambda^{opt} < \lambda^{pc}$ effects are opposite, and λ^{opt} is independent of η.

To implement a slower (faster) transition, the relative price the consumer faces should be below (above) the long run p_i^R. Let $\tilde{\tau}_i^{R,\lambda}(t)$ be the value of $\tilde{\tau}_i^R(t)$ required to implement a given λ. Then using (22) and (24), I find that under perfect competition:

$$\tilde{\tau}_i^{R,\lambda}(t) = \frac{\theta - 1 + \lambda}{\eta} \tilde{h}_i^R(t). \tag{31}$$

This tax rule reflects the result in Proposition 2: whenever $\gamma > \theta$, $\lambda^{opt} < 1 - \theta$ (see Lemma 4), and we find a negative relationship between the $\tilde{\tau}_i^{R,\lambda}(t)$ that implements welfare-maximizing consumption and $\tilde{h}_i^R(t)$. Conversely, $(\theta - 1 + \lambda^{opt})/\eta > 0$ if $\gamma < \theta$. Finally, from 23 we know $\tilde{h}_i^R(t)$ declines over time, converging to zero in the long run.

As described in Section 3.2.2, along the transition, strategic behavior by the monopolist increases the relative price p_i^R in excess of the increase in relative marginal costs. As a consequence, compared to the benchmark with $p_i^R = p_i^{R*}$, the shift in consumption from good i to b is already faster to begin with. One would thus expect that a subsidy on i, which slows down the transition, is more likely required to implement the optimal transition in the presence of market power. This is indeed the case:

Proposition 3. *Suppose goods are produced by monopolists. If* $\max\{\gamma, \theta\} > 0$*, the laissez-faire transition to the steady state is suboptimally fast. The optimal adjustment path can then be implemented by introducing a positive and declining subsidy on good i whenever* $\tilde{h}_i^R(0) > 0$ *and a positive and declining tax on good i if* $\tilde{h}_i^R(0) < 0$*. If* $\gamma = \theta = 0$*, the laissez-faire transition to the steady state is optimal, and any constant subsidy (tax) implements this path.*

Proof. See Appendix A.2.6 □

With habit formation (i.e. either γ or $\theta > 0$), the monopolist always implements a transition that is too rapid from a welfare perspective. First, the monopolist does not take into account the benefits of a slow transition in bringing down the aggregate habit H. Yet even if $\gamma = 0$, i.e. even if the benchmark habit plays no role in determining utility from consumption C, a welfare-maximizing policy slows down the shift in consumption from i to b implemented by the monopolist. This is because of the following. We know that along the transition, there is a benefit to quickly 'rebalancing' the consumption weights w_i such that they become more in line with actual consumption. The monopolist recognizes this too; as demand for good i falls over time, investing in the habit

[32] The rate at which consumption and habits adjust to the new steady state, $\xi\lambda^{opt}$, is always increasing in ξ. See Appendix A.3.2 for a proof.

becomes less valuable. As a response, the monopolist increases its markup to quickly divest habit and thus reduce the consumption weight. The monopolist however, does not take into account that this increases the consumption weight of all other goods, leading to a rapid shift in the w_i. From a welfare-perspective, this shift is too rapid. Hence, (partially) countering the monopolist's response to increase prices when habits are 'too high' increases welfare.

I can again solve for $\tilde{\tau}_i^{R,\lambda}(t)$, now under monopolistic supply. Equations (22) and (25) give

$$\tilde{\tau}_i^{R,\lambda}(t) = \frac{1}{\eta}\left[\frac{(\theta-1)\left[(\rho+\xi)+\xi\frac{\theta}{\eta-1}\right]+\lambda(\rho+\xi\lambda)}{(\rho+\xi)+\xi\frac{\theta}{\eta-1}}\right]\frac{\rho+\xi}{\rho+\xi(1+\lambda)}\tilde{h}_i^R(t), \tag{32}$$

where by definition, the bracketed term is zero for $\lambda = \lambda^{mc}$, and negative for $\lambda < \lambda^{mc}$. As by Lemma 4, $\lambda^{opt} < \lambda^{mc}$, we then find a negative relationship between the $\tilde{\tau}_i^{R,\lambda}(t)$ that implements welfare-maximizing consumption and $\tilde{h}_i^R(t)$.

5 Application

To illustrate the adjustment path of consumption and assess the potential quantitative implications of habit formation I evaluate the effects of an unanticipated 10 percent increase in the production cost of a subset of goods. Though insights apply more generally, for the sake of exposition, I interpret this cost shock in the context of food taxes, where the subset of goods are 'unhealthy foods', and the cost increase may resemble the introduction of a 10 percent charge on the saturated fat and sugars.[33,34] This charge may be fully passed through to consumers, or producers may act strategically and adjust markups in response to the levy. In either case, due to habits in consumption, demand for fatty and sugary foods will not instantly jump to the new steady state with a lower consumption of unhealthy foods. I compute the paths of consumption and prices as the economy converges to its new steady state. I compare the paths under perfect competition and monopolistic supply to the first-best consumption path. Temporary subsidies will be required to implement the optimal consumption path. These subsidies can be interpreted as temporary discounts on the permanent charge, or simply a slow phase-in of the charge. Finally, I compute the welfare gain of implementing an optimal path instead of the alternative adjustment paths.

To obtain numerical results I discretize the model. Further details about the discrete-time model setup can be found in Appendix C.

[33]This 10 percent charge is within the same order of magnitude as the fat tax introduced in Denmark in 2012 (Jensen and Smed, 2013). Note that I do not aim to perform a detailed policy simulation such as Allais et al. (2010). Given the stylized nature of the framework, the results should be primarily viewed as an indication of the quantitative significance of the habit internality.

[34]To ensure a clear distinction between the 'habit' tax (or subsidy) and the 'fat and sugar' tax, I will refer to the latter as a charge.

5.1 Parameter choices

The parameter values are determined as follows. Across the US and Europe, spending on food and nonalcoholic beverages accounts for about 12 percent of household spending (Eurostat, 2016; BLS, 2017). Of this 12 percent, about a third can be classified as 'unhealthy' (Mytton et al., 2007). Based on this, I set n, which I define as the share of goods subject to the charge, equal to 0.04. I separate the gross tax τ_i into two parts, the constant charge τ_{iC}, and the (potentially) time-varying 'habit tax' τ_{iH}, such that $\tau_i = \tau_{iC}\tau_{iH}$. Initial τ_{iC} and τ_{iH} are equal to 1 for all goods. Then, as of time $t = 0$, unhealthy food will be subject to an additional charge, such that for producers of goods $i \in [0, n]$, τ_{iC} increases to 1.1.

In the framework, the elasticity of substitution directly determines the short run price elasticity of demand (see (8)). Empirical estimates for the latter for specific consumer goods, including food categories, typically deliver low values, often below 1, suggesting complementarity (see for instance Andreyeva et al., 2010; Zhen et al., 2011; Green et al., 2013). Macro-level calibrations require values above 4 to match observed markups (Ravn et al 2006; 2010). For the main part of the numerical exercise I take a middle ground and set $\eta = 2$; I perform sensitivity analysis for $\eta = 0.4$ and $\eta = 1.2$.[35]

The habit parameters θ and γ are major determinants of the rate at which consumption transitions take place, and the policy required to maximize welfare along a transition. Several approaches can be used to infer the appropriate values for these parameters.

For θ, I consider empirical research on good-specific consumption persistence, and research that estimates both the short- and long run price elasticities of demand. Under the former approach, estimates for θ range from zero to 0.72, with a central value of about 0.3 (Carrasco et al., 2005; Zhen et al., 2011; Bronnenberg et al., 2012; Verhelst and Van den Poel, 2014). With the exception of Bronnenberg et al. (2012), these estimates use (a measure of) previous month or quarter consumption expenditure as a benchmark. The appropriate benchmark is however not immediate, and if habits are persistent, these estimates may either under- or over-estimate the 'real' θ. Bronnenberg et al. (2012) instead use geographic variation in brand preferences to elicit the causal effect of past experiences on future preferences. They find that 60 percent of the gap in brand preferences can be attributed to supply-side factors, while endogenous and persistent brand preferences explain 40 percent of the geographic variation in brand market shares. One can show this corresponds to $\theta = 0.4$.

An alternative estimation procedure for θ does not face the 'benchmarking' problem either. This approach is based on short- and long-run price elasticities of demand. From Section 3.1, I know these are equal to $\varepsilon_p^{SR} = \eta$ and $\varepsilon_p^{LR} = \eta/(1-\theta)$ respectively. Then $\theta = 1 - \varepsilon_p^{SR}/\varepsilon_p^{LR}$. Espey

[35] Since the analysis for the monopolist requires $\eta > 1$, I only evaluate perfect competition and first-best equilibria for $\eta = 0.4$.

et al. (1997) conduct a meta analysis of price elasticities for residential water consumption. Based on their median estimates for short and long run elasticities, I find θ equal to 0.41. Scott (2015) presents an overview of estimates of the elasticity of gasoline demand. The central value for θ based on these estimates is 0.6. Baltagi et al. (2000) estimate cigarette demand and also arrive at a value of 0.6.[36] Demand persistence for gasoline and cigarettes however likely overestimates the persistence of a 'representative' good: cigarettes are highly addictive and short-run gasoline demand is to a large extent determined by the vehicle a consumer owns. For this reason, I consider the estimate of 0.6 to be an upper bound for the appropriate θ and set $\theta = 0.4$. I perform sensitivity analysis for $\theta = 0.2$ and $\theta = 0.6$,

For γ, I consider estimates based on empirical evidence related to the Easterlin paradox and hedonic adaptation, aggregate consumption persistence and calibrations. High values for γ (close to 1) are required to explain the Easterlin paradox (Easterlin, 1974; Easterlin et al., 2010). Although evidence for happiness, or hedonic, adaptation is robust, the strong form of the Easterlin paradox, where long-run happiness is unaffected by income changes, is heavily contested (Clark, 1999; Oswald and Powdthavee, 2008; Stevenson and Wolfers, 2008; Easterlin et al., 2010). With incomplete adaptation, the value of γ is not easily determined, as reported happiness scores cannot be directly translated to utility units.

In my framework, to focus on consumption shifts across sectors, I abstract from saving and capital accumulation. If intertemporal consumption tradeoffs are taken into account, the aggregate habit parameter γ plays an additional role in determining the degree of aggregate consumption persistence.[37] Empirically estimating this persistence, Ravina (2005) and Alvarez-Cuadrado et al. (2012) find that a 1 percent increase in past aggregate consumption increases current consumption by 0.3 to 0.5 percent.[38] The corresponding estimate for γ then depends on the σ chosen. For $\sigma = 2$, γ lies in between 0.6 and 1, and higher γ are found for lower σ.[39]

Finally, I turn to calibrations. In a model that allows for saving, Abel (1990) requires values for γ close to 1 to explain the equity premium puzzle. Fuhrer (2000) introduces habits in a monetary policy model and estimates γ to fit the data. He arrives at a value of 0.8 to 0.9. Overall, evidence seems to suggest higher values for γ than θ. I follow Fuhrer (2000), set $\gamma = 0.8$, and perform sensitivity analysis for $\gamma = 0.4$ and $\gamma = 0.6$.

[36] Baltagi et al. (2000) compare a large number of models. I use the estimate of the model they consider best-performing.

[37] See for instance Carroll et al. (2000), Fuhrer (2000), Diaz et al. (2003) and Alvarez-Cuadrado et al. (2004). See also footnote 29 and Section 6.

[38] Dynan (2000) and Guariglia and Rossi (2002) estimate consumption persistence based on aggregate food consumption. As food is still a broad aggregate, I cannot readily reinterpret their estimates as estimates of θ or γ. They both find no or negative consumption persistence. However, their estimates, as well as those by Ravina (2005) and Alvarez-Cuadrado et al. (2012), suffer from the same 'benchmarking' problem discussed before.

[39] For $\sigma = 1$, aggregate consumption demand is independent of the habit. For $\sigma < 1$, $\gamma < 0$ is required to generate aggregate persistence.

Less empirical guidance exists regarding the speed of habit adjustment, ξ. Ravn et al. (2012) and Bronnenberg et al. (2012) find habits adjust very slowly over time; on an annual basis ξ is equal to 0.05 and 0.025 respectively. This slow adjustment is in line with Logan and Rhode (2010) and Atkin (2013), who find that prices (more than) 10 years in the past can partly explain current patterns of food consumption. This contrasts much of the literature which takes habits as equal to past-year consumption, as well as Carroll et al. (2000) who adopt an annual value of 0.2, and Constantinides (1990) who requires values as high as 0.6 to explain the equity premium puzzle. I take a 50 percent annual adjustment. As I estimate the model on a monthly basis ($dt = 1$ month), this gives $\xi = 0.06$. In the sensitivity analysis I consider $\xi = 0.03$ and $\xi = 0.12$. Finally, I set the monthly discount rate $\rho = 0.0035$,[40] the elasticity of marginal utility $\sigma = 1.5$, unit labor requirement $\delta_i = 1$ for all i, and total labor supply $L = 1$. With initial marginal production cost δ_i equal to 1 for all i, this gives initial steady state consumption, habits and prices equal to 1 for all goods. An overview of all baseline parameter values can be found in Table 1.

Table 1: Baseline parameter values

Parameter	Baseline values	Description
n	0.04	Share of goods subject to the charge
δ_i	1	Unit labor requirement
η	2	Elasticity of substitution
θ	0.4	Deep habit strength
γ	0.8	Aggregate habit strength
ξ	0.06	Habit adjustment speed (monthly)
ρ	0.0035	Rate of time preference (monthly)
σ	1.5	Elasticity of marginal utility
L	1	Labor supply

5.2 Results

Figure 1 shows the response of unhealthy food consumption and prices relative to 'non-unhealthy food' consumption following the introduction of the permanent charge on saturated fats and sugar at time 0. The dashed and dotted curves depict the response when goods are produced competitively or by monopolists respectively, without any additional policy intervention. The solid curves depict the optimal paths.

Under perfect competition, consumers face a one-off increase in prices (see Figure 1b). In response to this price increase, consumers instantly reduce relative consumption by 17 percent.[41]

[40] This corresponds to an annual discount rate of about 4 percent.

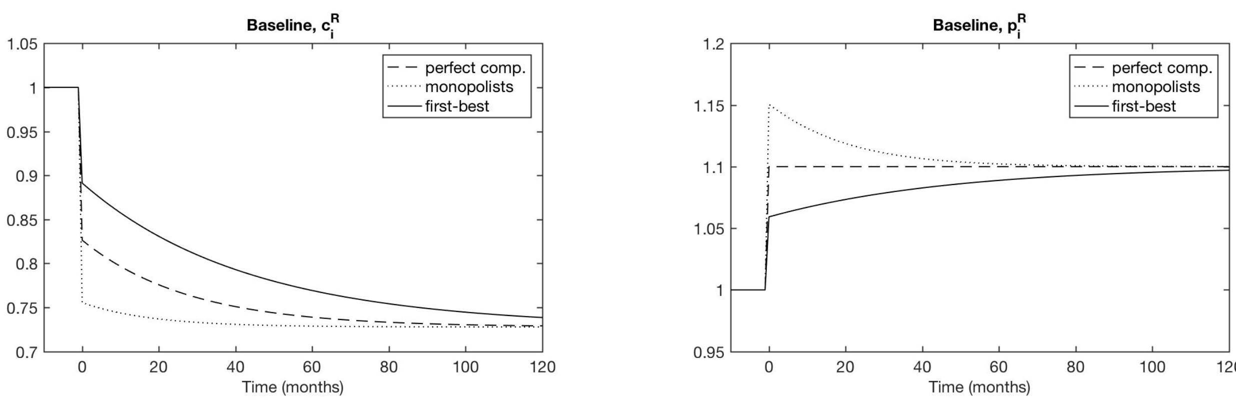
[41] Figure 1a depicts relative consumption paths. As the price change only affects 4 percent of goods, demand for all other goods increases by less than 1 percent at any point during the transition. The difference between changes

An additional 10 percent reduction is achieved as habits fall over time and consumption follows this drop in habits. As expected, the shift away from unhealthy foods is faster under monopolists: at $t = 0$ consumption immediately drops by 24 percent. Following this drop, habits quickly adjust. The rapid consumption response is the consequence of strategic behavior; monopolists increase prices for unhealthy foods relative to other goods by an additional 5 percent (see Figure 1b).

Along the transition to the new steady state, neither of the two paths described above are optimal. From Proposition 3 we know that the monopolist always implements a transition that is suboptimally fast. For our parameter values we have $\gamma > \theta$. Proposition 2 then informs us that also under perfect competition, the shift away from unhealthy food consumption under perfect competition is faster than optimal. Figure 1a confirms this: c_i^R is higher along the optimal path (solid curve) than along the paths where the transition is not specifically managed (dashed and dotted curves). In the optimum, time 0 unhealthy food consumption falls by only 11 percent. Consumption continues to drop afterwards, yet it takes more than 10 years until the full transition is accomplished. To ensure consumers select these first-best consumption levels, the relative price for unhealthy food should increase by only 5.9 percent initially, and then slowly rise to its long run level of 1.1.

Figure 1: Relative consumption and prices along the transition

(a) (b)



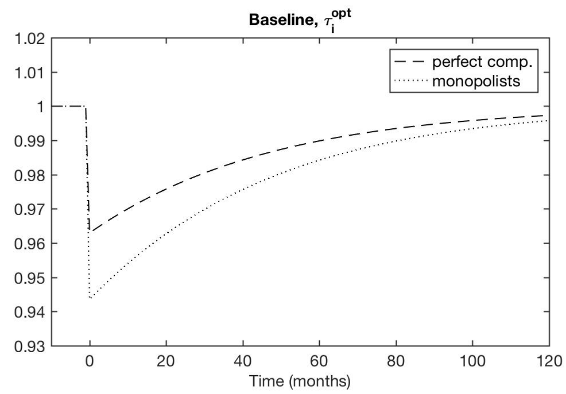
The curves depict responses to the unanticipated introduction of the 10 percent charge at $t = 0$, where I assume the economy is in steady state for all $t < 0$. Responses are shown for a good $i \in [0, n]$ relative to any good $b \in (n, 1]$.

To implement the price path depicted by the solid line in Figure 1b we require a temporary subsidy to unhealthy food (i.e. discount to the food charge). Under perfect competition, this would amount to a subsidy of about 3.4 percent (gross tax of 0.966, see Figure 2). This is about a third

in absolute unhealthy food consumption, and consumption of unhealthy food relative to non-unhealthy food goods is thus small, and I use the two concepts interchangeably.

of the total 10 percent charge. A larger transitory subsidy is required if goods are produced by monopolists, who initially increase prices in excess of the charge. Now, an initial subsidy of 5.6 percent, which falls to 4.6 percent after 1 year and 1.6 percent after 5 years implements the optimal adjustment path.

Figure 2: Relative taxes to implement first-best transition



The curves depict the path of taxes required to implement the first-best adjustment path, in response to the unanticipated introduction of the 10 percent charge at $t = 0$, where I assume the economy is in steady state for all $t < 0$. Taxes are shown for a good $i \in [0, n]$ relative to any good $b \in (n, 1]$.

The above-mentioned subsidies have a large impact on prices and consumption choices as the economy reduces its unhealthy food consumption. This raises the question of whether this policy also generates sizable welfare gains. For this purpose, I compute the consumption-equivalent welfare loss due to the transition.[42] Here I take into account that the charge may be implemented as a welfare-improving policy to begin with, and may thus not only set in motion a transition to a new steady state, but also move the economy away from a *distorted* steady state.[43] To separate these two effects, I compute two losses, one which solely captures the loss due to the transition, and one which, in addition, captures the benefit from correcting a previously uncorrected externality. Results are presented in Table 2.[44]

From Figure 1a I know that consumption adjustment is relatively slow along the optimal path, and fast under monopolistic supply. Hence, I expect welfare to be highest (losses to be lowest) in the first-best transition, followed by the transition under perfect competition and monopolists. This is confirmed by Table 2. Table 2 reports the consumption-equivalent welfare loss of the transition

[42] More formally, let $W_X(C_X)$ be welfare under consumption path X and $W^*(C^*)$ welfare if the economy is in steady-state. Then the (steady-state) consumption-equivalent welfare loss is β_X, with β_X implicitly determined by $W_X(C_X) = W^*((1-\beta_X)C^*)$.

[43] This can for instance be motivated by the existence of publicly-paid healthcare systems, which cause consumers to not carry the full burden of dietary choices.

[44] See Appendix C.1 for more details.

for four cases; the three cases considered in the paper, and the case where all c_i are set to their respective steady-state immediately ('no transition', bottom row).[45] Reported losses are small, about 0.008%; as the 10% long-run price shock only affects a share $n = 0.04$ of goods this is not surprising. Optimally managing the transition reduces welfare losses by 0.0056% to 0.0012%; this is 1.4%-6.9% of the loss under first-best. This reduction in welfare losses is very similar when I take into account benefits from correcting for the production externality.

Table 2: Consumption-equivalent loss of transition

Case	Loss excluding 'externality benefit'	Loss including 'externality benefit'
First-best	0.0811%	-0.0067%
Perfect competition	0.0823%	-0.0055%
Monopolists	0.0851%	-0.0027%
No transition	0.0867%	-0.0014%

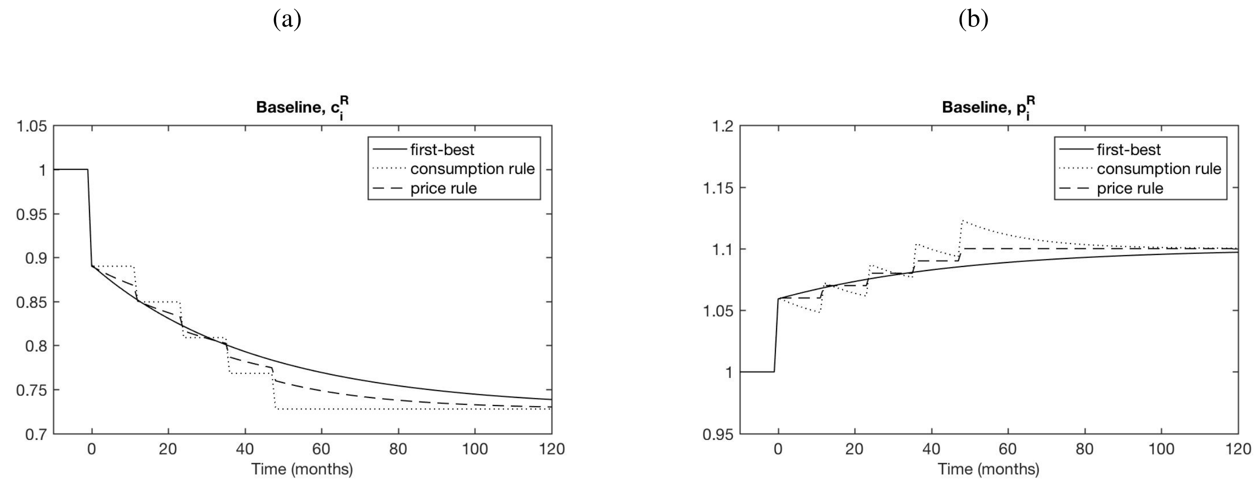
Rule of thumb policy Even though the first-best adjustment path minimizes welfare losses, the other two paths have a clear advantage in terms of implementation. Under the perfect competition and monopolists path, policy takes the form of a one-off increase in the 'unhealthy-food' charge, absent any further intervention. Of these two paths, the perfect competition path, where consumers face a flat price schedule, performs best. In this subsection, I consider two 'rule of thumb' policies which aim to reduce welfare losses compared to this perfect competition path, yet are more straightforward to implement than the optimal path. The first rule of thumb policy targets consumption; it is a 'unhealthy food' quota, that is lowered each year. In the first year, the rule imposes a 11 percent drop in c_i^R. Then, each year, for four consecutive years, the c_i^R is lowered by an equal amount, such that after 4 years, it reaches its steady state.[46] The second rule of thumb policy targets prices. It sets a relative price of 'unhealthy food' equal to 1.06 in the first year, and increases this price by 0.01 point each year for four years thereafter. Figure 3 presents the paths of c_i^R and p_i^R under both rules of thumb, and in the first-best transition. As is clear from Figure 3a, both rules implement a shift that is somewhat slower than first-best initially, yet reaches the steady state sooner. Under the consumption rule, (implicit) prices overshoot the long run equilibrium for a substantial period of time. The simple policy rules do not always improve upon the allocation with the flat price schedule (Table 2, second column). Under the consumption rule, the consumption-

[45]Note that even though consumption immediately jumps to its long run level, habits still require time to adjust.

[46]The 4 year period is chosen as follows. For the parameter values in Table 1, $\lambda^{opt} = 0.3656$ (see Appendix A.3.1). The adjustment speed, $\xi * \lambda$ is then about 2.2 percent a month. A rough approximation of the total adjustment period in turn gives $1/0.022 = 45.5$ months≈4 years. The 11 percent initial drop in c_i^R is set close to the initial drop in the optimal path (10.87%). The approach to determine the rule of thumb price path is equivalent, with the initial 6% price increase set close to the initial increase in first-best (5.92%).

equivalent loss of the transition is 0.0824%, i.e. it performs better than the monopolist (and no transition) path, but is just shy of beating the perfect competition case. The price rule performs better, here I find a loss of 0.0818%.[47]

Figure 3: Relative consumption and prices under rules of thumb



Curves depict policy rules where policy is introduced at $t = 0$. I assume the economy is in steady state for all $t < 0$. Paths are shown for a good $i \in [0,n]$ relative to any good $b \in (n,1]$.

Sensitivity Results for alternative parameter values are presented in Figures D.1-D.3 in Appendix D. The effects of a change in the value of a particular parameter are fully in line with the results presented in Sections 3-4 and Lemmas 1-5. A lower short run price elasticity of demand, η, reduces the short- and long-run consumption responses in response to the charge. Interestingly, for $\eta = 1.2$, consumption undershoots its long run equilibrium if goods are produced by monopolists. This is in line with Lemma 3, which states that $\lambda^{mc} > 1$ if $\theta > 0$ and $\eta - 1 < (1-\theta)\xi/(\rho+\xi)$.[48] The long run price elasticity is also determined by θ, with higher values for θ implying larger long-run consumption adjustment in response to the charge. A lower θ (good-specific demand less dependent of habits) and higher ξ (faster habit adjustment) clearly increase the rate at which the economy transitions to the steady-state. Proposition 3 stated that in the log-linearized model, the perfection competition transition equals first-best when $\gamma = \theta$. I consider this case when I set $\gamma = 0.4$ (Figures D.1-D.3, third row, left column), and find that also here, the approximation is very accurate; the respective c_i^R deviate by less than 0.003.

[47] When I take into account the benefit from correcting the externality, as in the rightmost column of Table 2, I find values of -0.0060% and -0.0066% respectively. Now, the consumption rule does improve welfare compared to the perfect competition path.

[48] For the parameter values in Table 1, $(1-\theta)\xi/(\rho+\xi) = 0.57$. Thus, this condition is satisfied whenever $\eta < 1.57$.

The result that habit formation might call for substantial discounts on the initial charge is not very sensitive to parameter choices. With the exception of the last case discussed above, optimal short-run discounts on the charge range from 20 to 60 percent. In addition, within the range I consider, the variation in $\tau_i^{opt}(0)$ is almost fully driven by variation in the two habit parameters, θ and γ.

Accuracy of the linear approximation As a final exercise, I compare the numerical results to the linear approximation of the consumption path in (21). I find that this approximation is accurate. For the parameter values in Table 1 we have $\lambda^{pc} = 0.6$, $\lambda^{mc} = 0.91$ and $\lambda^{opt} = 0.37$ and $\tilde{h}_i^R(0) = 0.37$. Then, from (21), time 0 unhealthy-food relative to non-unhealthy-food consumption should equal 0.84, 0.75, and 0.90 under perfect competition, monopolists and the optimal path respectively. Comparing these values to the results discussed above I find that the approximation is off by at most 0.01 point. The linear approximation is accurate too regarding the adjustment speed. For the change in c_i^R that remains after the initial drop, the approximation predicts a half life of 19, 13 and 32 months for perfect competition, monopolists and the optimal path respectively.[49] Comparing these half-lives to those in Figure 1a reveals a bias of at most one month.

6 Discussion

Though rich in several aspects, the modeling framework remains highly stylized, with several assumptions made to facilitate the analysis and exposition. Below, I briefly discuss several of such assumptions, address the sensitivity of results to these assumptions, and how the framework can be flexibly adjusted to account for them.

Labor-leisure and consumption-savings decision Models of aggregate habit formation conventionally feature either a labor-leisure (e.g. Cremer et al., 2010), or consumption-savings decision (e.g. Carroll et al., 2000). These features allow for aggregate consumption persistence, i.e. an adjustment in labor supply or savings, and hence aggregate consumption, in response to aggregate habits. The focus of the analysis in this paper is on the within-bundle allocation and persistence. Here, one can show that the within-bundle (relative) allocation c_i^R is independent of the aggregate consumption level (see (8)). This implies that all results that concern the within-bundle allocation continue to hold if capital accumulation would play a role, or labor supply would be endogenous.[50] Even though results regarding c_i^R, and also p_i^R and optimal τ_i^R are robust to this extension, results for consumption, price and tax *levels* might be slightly adjusted. For instance, as explained in

[49] For the approximation, the half life T is computed by solving $e^{-\xi\lambda T} = 0.5$.

[50] Details are available on request.

footnote 29, implementation of optimal steady-state consumption would require a positive tax that is common across goods and corrects for the aggregate consumption internality. As following a cost shock to a small subset of goods, aggregate consumption and habits change little compared to relative consumption and habits, I expect changes in the optimal tax level to be of second-order to relative tax changes.[51]

Partial projection bias Thoughout the paper I have maintained the assumption that there is no internalization of the habit formation process on part of the consumer (see Assumption 1). This assumption is arguably strong; a setup where the consumer internalizes the effect of current consumption on future habits to at least some degree might be more realistic. As long as goods are produced under perfect competition, the model can be straightforwardly adjusted to such imperfect internalization. One can then show that equilibrium consumption with imperfect habit internalization falls in between the equilibria with no habit internalization and optimal consumption, implying that Propositions 1 and 2 will continue to hold.[52] The equilibrium for monopolistic supply is much more complex to solve as both current demand and prices become a function of future prices, and involves firms facing a time-inconsistency problem (see Nakamura and Steinsson, 2011); I leave the evaluation of within-bundle transitions for this case to future research.

Heterogeneous habits For certain types of consumer demand, habits may be more important, or slower to adjust. Goods for instance may have addictive properties that justify a greater elasticity of demand with respect to the habit θ, or a lower habit adjustment parameter ξ. Though the framework can be easily adjusted to allow for good-specific θ_i and ξ_i, the transition cannot be approximated through a unique λ; any detailed analysis will have to rely on numerics instead. As the intuition behind the roles of these two parameters is independent of the specific levels of θ and γ, I expect that general insights will extend to the case of heterogeneous habits; consumption transitions are likely inefficiently slow for high θ_i, and faster the greater ξ_i.

Source of cost shock Often, a consumption transition is initiated by policy aimed at correcting previously-existing market failures, for instance fat or environmental taxes. In Appendix B I extend the model with a production externality (e.g. an environmental externality with production damages). Here, I show that, if the marginal external effect is independent of the level of use, all results concerning the optimal rate of relative consumption adjustment carry through. The optimal corrective tax $\tilde{\tau}_i^R$ can then be decomposed in a time-invariant charge that corrects the production externality, and the time-variant tax as in (31) and (32). An interesting extension for future research

[51] For instance, for the numerical example in Section 5 I find that the shift in c_i^R relative to h_i^R is at least a factor 25 larger than the shift in C relative to H.

[52] Detailed derivations are available on request.

would be to consider the case where the marginal external effect is a function of the consumption level, or cumulative use is restricted. In such cases, habit formation, even if perfectly internalized, would likely lead to optimal corrective taxes that are time-varying.

7 Conclusion

This paper studies consumption choices when consumption is subject to good-specific habit formation. I develop a stylized representative-consumer model where I explicitly distinguish between two roles of habits. First, good-specific habits cause the allocation of consumption across goods to be persistent; shifts within the consumption bundle are slow. Second, these habits jointly determine the benchmark against which consumption is evaluated; the higher habits, the higher the 'standard of living' the consumer is accustomed to and the lower is welfare for a given level of consumption. I consider the case where due to projection bias, the consumer does not internalize the fact that current consumption affects future habits, and characterize the equilibrium consumption choices and prices when goods are produced under perfect competition or by forward-looking monopolists.

I find that the steady-state allocation of consumption across goods is independent of market power. As all goods are to an equal extent subject to habit formation, projection bias does not distort the steady-state consumption allocation. The transition toward the steady state may however be suboptimal. Whenever goods are produced by monopolists, strategic pricing speeds up shifts within the consumption bundle, and cause those to be suboptimally fast. In this case, the optimal transition can be implemented by a tax that slows down the transition to the steady state. In the absence of market power, the consumer adjusts its consumption bundle suboptimally fast if the second, 'welfare', role of habits is particularly strong. If instead the persistence effect of habits dominates, a tax policy that speeds up the transition is preferred.

While the model and its discussion are stylized, its applications are numerous. In the upcoming decades, some major shifts will likely occur in our consumption patterns. Increased water shortages in many regions in the world necessitate consumers to reduce water use. Resource scarcity and concerns about climate change will call for a reduction in energy use, especially if the cost of renewable energy remains high, and in many countries, taxes on fatty or sugary foods to induce consumers to adopt a healthier diet are currently on the table. Such (policy-induced) shifts in our consumption bundles will not happen instantaneously, with habit formation being one reason for such slow transitions. In all these instances, a relevant question is whether there is a role for policy in managing the rate at which such change occurs. For instance, in the context of fat taxes, it is optimal to force consumers to quickly get rid of unhealthy dietary habits by introducing hefty initial rates? Or is it perhaps preferred to allow consumers to slowly adjust their demand for fatty foods, by implementing a tax that starts low and increases over time?

In a numerical exercise I apply the model to answer the latter question. More specifically, I consider the introduction of a 10 percent charge on a subset of goods ('unhealthy food'), which induces a transition away from these goods. This application reveals that the theoretical effects are also quantitatively meaningful. I find that the transition under perfect competition and monopolistic supply is substantially faster than the first-best transition; the first-best path calls for a lower immediate reduction in 'unhealthy food' consumption, and allows the remaining reduction to slowly materialize over time. To implement this path, the policymaker can offers consumers an initial discount of as much as 60 percent of the long-run charge. The optimal path requires careful management of consumption and/or prices. I evaluate two rule of thumb policies that are easier to implement and have the potential to bring welfare closer to first-best compared to the one-off charge.

In the above examples the shift in the consumption bundle is policy-induced to begin with. The same questions and insights however apply if the cause of the shift is external. Consider for example the common call for policy action when gasoline prices increase due to shifts on world oil markets. Also shifts in food prices, caused by misharvests or increased openness to trade, are often followed by appeals for government intervention such as (temporary) subsidies or tax breaks.[53] The framework and numerical results in this paper provide support for such measures; with habit formation and projection bias, a policy that allows people to partly postpone adjustment in consumption is welfare-improving.

Acknowledgments

I wish to in particular express my gratitute to Reyer Gerlagh and Sjak Smulders for in-depth discussions. I would also like to thank Lint Barrage, Davide Cerruti, Patricio Dalton, Niko Jaakkola, Olof Johansson-Stenman, Ramón López, Rick van der Ploeg, Per Krusell, Mauricio Rodriguez Acosta, Daniel Spiro, Joe Vecci, Rob Williams and Sophie Lian Zhou for helpful discussions and comments. I thank participants at various research seminars and conferences for helpful suggestions and comments.

References

Abel, Andrew B (1990) "Asset prices under habit formation and catching up with the Joneses," *The American Economic Review*, pp. 38–42.

[53]The latter example relates to the work by Atkin (2013), who documents that habits reduce the nutritional gains from trade in India, as consumers continue to favor foods that were relatively inexpensive in the past.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous (2012) "The Environment and Directed Technical Change," *The American Economic Review*, pp. 131–166.

Allais, Olivier, Patrice Bertail, and Véronique Nichèle (2010) "The effects of a fat tax on French households' purchases: a nutritional approach," *American Journal of Agricultural Economics*, Vol. 92, pp. 228–245.

Alonso-Carrera, Jaime, Jordi Caballé, and Xavier Raurich (2005) "Growth, habit formation, and catching-up with the Joneses," *European Economic Review*, Vol. 49, pp. 1665–1691.

Alvarez-Cuadrado, Francisco, Jose Maria Casado, Jose Maria Labeaga, and Dhanoos Sutthiphisal (2012) "Envy and habits: panel data estimates of interdependent preferences," Banco de Espana Working Papers 1213, Banco de Espana.

Alvarez-Cuadrado, Francisco, Goncalo Monteiro, and Stephen J Turnovsky (2004) "Habit formation, catching up with the Joneses, and economic growth," *Journal of Economic Growth*, Vol. 9, pp. 47–80.

Andreyeva, Tatiana, Michael W Long, and Kelly D Brownell (2010) "The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food," *American Journal of Public Health*, Vol. 100, pp. 216–222.

Aronsson, Thomas and Olof Johansson-Stenman (2014) "Positional preferences in time and space: Optimal income taxation with dynamic social comparisons," *Journal of Economic Behavior & Organization*, Vol. 101, pp. 1–23.

Atkin, David (2013) "Trade, Tastes, and Nutrition in India," *The American Economic Review*, Vol. 103, pp. 1629–63.

Baltagi, Badi H, James M Griffin, and Weiwen Xiong (2000) "To pool or not to pool: Homogeneous versus heterogeneous estimators applied to cigarette demand," *Review of Economics and Statistics*, Vol. 82, pp. 117–126.

Bernheim, B Douglas and Antonio Rangel (2007) "Behavioral public economics: Welfare and policy analysis with non-standard decision-makers," in Hannu Vartianen Peter Diamond ed. *Behavioral Economics and its Applications*, pp. 7–77: Princeton University Press.

BLS (2017) "Consumer Expenditures in 2015," Accessed April 18, 2017.

Bronnenberg, Bart J, Jean-Pierre H Dubé, and Matthew Gentzkow (2012) "The Evolution of Brand Preferences: Evidence from Consumer Migration," *The American Economic Review*, Vol. 102, pp. 2472–2508.

Campbell, John Y and John H Cochrane (1999) "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, Vol. 107.

Carrasco, Raquel, Jose M Labeaga, and J David López-Salido (2005) "Consumption and Habits: Evidence from Panel Data*," *The Economic Journal*, Vol. 115, pp. 144–165.

Carroll, Christopher D (2000) "Solving consumption models with multiplicative habits," *Economics Letters*, Vol. 68, pp. 67–77.

Carroll, Christopher D, Jody Overland, and David N Weil (2000) "Saving and growth with habit formation," *The American Economic Review*, pp. 341–355.

Choudhary, M. Ali, Paul Levine, Peter McAdam, and Peter Welz (2012) "The happiness puzzle: analytical aspects of the Easterlin paradox," *Oxford Economic Papers*, Vol. 64, pp. 27–42.

Clark, Andrew E (1999) "Are wages habit-forming? Evidence from micro data," *Journal of Economic Behavior & Organization*, Vol. 39, pp. 179–200.

Conlin, Michael, Ted O'Donoghue, and Timothy J Vogelsang (2007) "Projection bias in catalog orders," *The American Economic Review*, pp. 1217–1249.

Constantinides, George M. (1990) "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, Vol. 98, pp. pp. 519–543.

Cremer, Helmuth, Philippe De Donder, Dario Maldonado, and Pierre Pestieau (2010) "Commodity taxation under habit formation and myopia," *The BE Journal of Economic Analysis & Policy*, Vol. 10.

Dalton, Patricio S and Sayantan Ghosal (2011) "Behavioral decisions and policy," *CESifo Economic Studies*, Vol. 57, pp. 560–580.

DellaVigna, Stefano (2009) "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, Vol. 47, pp. 315–72.

Diaz, Antonia, Josep Pijoan-Mas, and José-Victor Rios-Rull (2003) "Precautionary savings and wealth distribution under habit formation preferences," *Journal of Monetary Economics*, Vol. 50, pp. 1257–1291.

Doi, Junko and Kazuo Mino (2008) "A variety-expansion model of growth with external habit formation," *Journal of Economic Dynamics and Control*, Vol. 32, pp. 3055–3083.

Dynan, Karen E (2000) "Habit formation in consumer preferences: Evidence from panel data," *The American Economic Review*, pp. 391–406.

Easterlin, Richard A (1974) "Does economic growth improve the human lot? Some empirical evidence," *Nations and households in economic growth*, Vol. 89, pp. 89–125.

Easterlin, Richard A, Laura Angelescu McVey, Malgorzata Switek, Onnicha Sawangfa, and Jacqueline Smith Zweig (2010) "The happiness-income paradox revisited," *Proceedings of the National Academy of Sciences*, Vol. 107, pp. 22463–22468.

Espey, Molly, James Espey, and W Douglass Shaw (1997) "Price elasticity of residential demand for water: A meta-analysis," *Water Resources Research*, Vol. 33, pp. 1369–1374.

Eurostat (2016) "Household consumption by purpose," Accessed 18 April 2017.

Frederick, Shane and George Loewenstein (1999) "Hedonic adaptation."

Fuhrer, Jeffrey C (2000) "Habit formation in consumption and its implications for monetary-policy models," *The American Economic Review*, pp. 367–390.

Gerlagh, Reyer, Snorre Kverndokk, and Knut Einar Rosendahl (2009) "Optimal timing of climate change policy: Interaction between carbon taxes and innovation externalities," *Environmental and Resource Economics*, Vol. 43, pp. 369–390.

Green, Rosemary, Laura Cornelsen, Alan D Dangour, Rachel Turner, Bhavani Shankar, Mario Mazzocchi, and Richard D Smith (2013) "The effect of rising food prices on food consumption: systematic review with meta-regression," *British Medical Journal*, Vol. 346, p. f3703.

Guariglia, Alessandra and Mariacristina Rossi (2002) "Consumption, habit formation, and precautionary saving: evidence from the British Household Panel Survey," *Oxford Economic Papers*, Vol. 54, pp. 1–19.

Jensen, Jørgen Dejgård and Sinne Smed (2013) "The Danish tax on saturated fat - short run effects on consumption, substitution patterns and consumer prices of fats," *Food Policy*, Vol. 42, pp. 18–31.

Koehne, Sebastian and Moritz Kuhn (2014) "Optimal taxation in a habit formation economy," *Journal of Public Economics*.

Layard, Richard (2006) "Happiness and Public Policy: a Challenge to the Profession*," *The Economic Journal*, Vol. 116, pp. C24–C33.

Ljungqvist, Lars and Harald Uhlig (2000) "Tax policy and aggregate demand management under catching up with the Joneses," *The American Economic Review*, pp. 356–366.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin (2000) "Projection bias in predicting future utility."

Logan, Trevon D and Paul W Rhode (2010) "Moveable Feasts: A New Approach to Endogenizing Tastes," *Manuscript: The Ohio State University*.

Monteiro, Goncalo, Adam Cook, and Sanjoy Dey (2013) "Optimal tax policy under habit formation and capital utilization," *Journal of Macroeconomics*, Vol. 37, pp. 230–248.

Mytton, Oliver, Alastair Gray, Mike Rayner, and Harry Rutter (2007) "Could targeted food taxes improve health?" *Journal of epidemiology and community health*, Vol. 61, pp. 689–694.

Nakamura, Emi and Jón Steinsson (2011) "Price setting in forward-looking customer markets," *Journal of Monetary Economics*, Vol. 58, pp. 220–233.

Oswald, Andrew J and Nattavudh Powdthavee (2008) "Does happiness adapt? A longitudinal study of disability with implications for economists and judges," *Journal of Public Economics*, Vol. 92, pp. 1061–1077.

Pollak, Robert A. (1970) "Habit Formation and Dynamic Demand Functions," *Journal of Political Economy*, Vol. 78, pp. pp. 745–763.

Rabin, M. (2002) "A Perspective of Psychology and Economics," *European Economic Review*, Vol. 46, pp. 657–685.

Ravina, Enrichetta (2005) "Habit persistence and keeping up with the Joneses: evidence from micro data."

Ravn, Morten O, Stephanie Schmitt-Grohe, and Martin Uribe (2010) "Incomplete cost pass-through under deep habits," *Review of Economic Dynamics*, Vol. 13, pp. 317–332.

Ravn, Morten O., Stephanie Schmitt-Grohé, and Martín Uribe (2012) "Consumption, government spending, and the real exchange rate," *Journal of Monetary Economics*, Vol. 59, pp. 215–234.

Ravn, Morten, Stephanie Schmitt-Grohé, and Martin Uribe (2006) "Deep habits," *The Review of Economic Studies*, Vol. 73, pp. 195–218.

Ryder, Harl E and Geoffrey M Heal (1973) "Optimal growth with intertemporally dependent preferences," *The Review of Economic Studies*, pp. 1–31.

Samson, Alain (2014) "The Behavioral Economics Guide."

Scott, K Rebecca (2015) "Demand and price uncertainty: Rational habits in international gasoline demand," *Energy*, Vol. 79, pp. 40–49.

Smed, Sinne (2012) "Financial penalties on foods: the fat tax in Denmark," *Nutrition Bulletin*, Vol. 37, pp. 142–147.

State of California (2015) "Storage in Major Reservoirs."

Stevenson, Betsey and Justin Wolfers (2008) "Economic growth and subjective well-being: Reassessing the Easterlin paradox,"Technical report, National Bureau of Economic Research.

The Guardian (2016) "World Health Organisation urges all countries to tax sugary drinks," Online, October 11.

The New York Times (2015) "Beneath California Crops, Groundwater Crisis Grows," April 5.

The Washington Post (2012) "What the world can learn from Denmark's failed fat tax," November 11.

Tirole, Jean (2002) "Rational irrationality: Some economics of self-management," *European Economic Review*, Vol. 46, pp. 633–655.

Turnovsky, Stephen J and Goncalo Monteiro (2007) "Consumption externalities, production externalities, and efficient capital accumulation under time non-separable preferences," *European Economic Review*, Vol. 51, pp. 479–504.

Verhelst, Benjamin and Dirk Van den Poel (2014) "Deep habits in consumption: A spatial panel analysis using scanner data," *Empirical Economics*, Vol. 47, pp. 959–976.

Zhen, Chen, Michael K Wohlgenant, Shawn Karns, and Phillip Kaufman (2011) "Habit formation and demand for sugar-sweetened beverages," *American Journal of Agricultural Economics*, Vol. 93, pp. 175–193.

A Derivations and proofs

A.1 Detailed derivations

A.1.1 Equations (11), (12) and (25)

The good i monopolist maximizes $V_i(t) = \int_t^{\infty} e^{-\int_t^{\nu} r(x)dx} c_i(\nu)\left[p_i(\nu) - \delta_i \tau_i(\nu)\right] d\nu$, by choosing the path of supply, $[c_i(\nu)]_{\nu=t}^{\nu=\infty}$, taking into account demand, (8), and the process of good-specific habit formation, (6). The producer takes as given the price index P and aggregate habit H. Hence it solves the following Hamiltonian:

$$\mathscr{H} = c_i \left[p_i - \delta_i p_L \tau_i\right] + \kappa_{p_i} \left[c_i^{-\frac{1}{\eta}} \left(\frac{h_i}{H}\right)^{\frac{\theta}{\eta}} C^{\frac{1}{\eta}} - p_i\right] + \kappa_{h_i} \left[\xi \left(c_i - h_i\right)\right], \tag{A.1}$$

where κ_{p_i} is the shadow value of inverse demand p_i and κ_{h_i} is the shadow value of habits h_i and I have already substituted (4) in (3). This gives the following FOC:

$$\begin{aligned} &[c_i] \quad p_i - \delta_i p_L \tau_i - \kappa_{p_i} \tfrac{1}{\eta} \tfrac{p_i}{c_i} + \xi \kappa_{h_i} = 0 \\ &[p_i] \qquad\qquad c_i - \kappa_{p_i} = 0 \\ &[h_i] \qquad \tfrac{\theta}{\eta} \kappa_{p_i} \tfrac{p_i}{h_i} - \xi \kappa_{h_i} = r \kappa_{h_i} - \dot{\kappa}_{h_i} \end{aligned} \tag{A.2}$$

Then I substitute $\kappa_{p_i} = c_i$ (see FOC with respect to p_i) in the FOCs for c_i and h_i. This gives (11) and (12). Next, I take the time derivative of (11):

$$\dot{\kappa}_{h_i} = -\frac{1}{\xi} \left[\frac{\eta - 1}{\eta} \dot{p}_i - \delta_i p_L \tau_i \left[\frac{\dot{p}_L}{p_L} + \frac{\dot{\tau}_i}{\tau_i}\right]\right], \tag{A.3}$$

which with the FOCs for h_i gives

$$\kappa_{h_i} = \frac{1}{r + \xi} \left[p_i \frac{\eta - 1}{\eta} \left[\frac{\theta}{\eta - 1} \frac{c_i}{h_i} - \frac{1}{\xi} \frac{\dot{p}_i}{p_i}\right] + \frac{1}{\xi} \delta_i p_L \tau_i \left[\frac{\dot{p}_L}{p_L} + \frac{\dot{\tau}_i}{\tau_i}\right]\right]. \tag{A.4}$$

I then substitute this result in (11) to find the following solution for p_i:

$$p_i = \delta_i p_L \tau_i \frac{\eta}{\eta - 1} \left[\frac{1 - \frac{1}{r+\xi} \left[\frac{\dot{p}_L}{p_L} + \frac{\dot{\tau}_i}{\tau_i}\right]}{1 + \frac{1}{\eta - 1} \frac{\xi}{r+\xi} \theta \frac{c_i}{h_i} - \frac{1}{r+\xi} \frac{\dot{p}_i}{p_i}}\right]. \tag{A.5}$$

This equation is in turn used to obtain the steady-state price, (15), and loglinearized to arrive at (25).

A.1.2 Equations (26), (27) and (30)

To maximize (1) subject to (2)-(7) by choosing $[c_i(\nu)]_{i\in[0,1],\nu=t}^{\nu=\infty}$ I write the following Hamiltonian:

$$\mathscr{H} = \frac{(CH^{-\gamma})^{1-\sigma}}{1-\sigma} + \mu_C\left[\left[\int_0^1 \left(\frac{h_i}{H}\right)^{\frac{\theta}{\eta}} c_i^{\frac{\eta-1}{\eta}} di\right]^{\frac{\eta}{\eta-1}} - C\right] + \mu_H\left[\left[\int_0^1 h_i^{\frac{\eta-1+\theta}{\eta}} di\right]^{\frac{\eta}{\eta-1+\theta}} - H\right]$$
$$+\mu_L\left[L - \int_0^1 \delta_i c_i di\right] + \mu_{h_i}\left[\xi\left(c_i - h_i\right)\right], \tag{A.6}$$

where μ_C, μ_H and μ_L are the shadow values of effective consumption, aggregate habit and labor respectively, μ_{h_i} is the shadow value of the good-specific habit, and (3) and (5) have been slightly rewritten by substituting in (4). This gives the following FOC:

$$\begin{aligned}
&[C] && (CH^{-\gamma})^{1-\sigma}\tfrac{1}{C} - \mu_C = 0 \\
&[H] && -\gamma(CH^{-\gamma})^{1-\sigma}\tfrac{1}{H} - \mu_C\tfrac{\theta}{\eta-1}\tfrac{C}{H} - \mu_H = 0 \\
&[c_i] && \mu_C C^{\frac{1}{\eta}}\left(\tfrac{h_i}{H}\right)^{\frac{\theta}{\eta}} c_i^{\frac{\eta-1}{\eta}}\tfrac{1}{c_i} - \delta_i\mu_L + \xi\mu_{h_i} = 0 \\
&[h_i] && \mu_C\tfrac{\theta}{\eta-1}\left(\tfrac{c_i}{C}\right)^{-\frac{1}{\eta}}\left(\tfrac{h_i}{H}\right)^{\frac{\theta}{\eta}}\tfrac{c_i}{h_i} + \mu_H\left(\tfrac{h_i}{H}\right)^{-\frac{1-\theta}{\eta}} - \xi\mu_{h_i} = \rho\mu_{h_i} - \dot{\mu}_{h_i}
\end{aligned} \tag{A.7}$$

Next define $\check{\mu}_L \equiv \mu_L/\mu_C$ and $\check{\mu}_{h_i} \equiv \mu_{h_i}/\mu_C$, such that $\dot{\check{\mu}}_{h_i} = \check{\mu}_{h_i}\left[\frac{\dot{\mu}_{h_i}}{\mu_{h_i}} - \frac{\dot{\mu}_C}{\mu_C}\right]$, and rewrite (8) to

$$p_i = \left(\frac{c_i}{C}\right)^{-\frac{1}{\eta}}\left(\frac{h_i}{H}\right)^{\frac{\theta}{\eta}}. \tag{A.8}$$

With the FOC for c_i this gives (26). Then to arrive at (27), I first substitute $\mu_C = (CH^{-\gamma})^{1-\sigma}C^{-1}$ in the FOC for H. This gives $\mu_H = -\mu_C\frac{C}{H}\left[\gamma + \frac{\theta}{\eta-1}\right]$. I then substitute these results for μ_C and μ_H and (A.8) in the FOC for h_i to find

$$\check{\mu}_{h_i} = \frac{1}{\check{r}+\xi}\left[\dot{\check{\mu}}_{h_i} + \frac{c_i}{h_i}p_i\left[\frac{\theta}{\eta-1} - \left[\gamma + \frac{\theta}{\eta-1}\right]\left(\frac{h_i}{H}\right)^{\frac{\eta-1}{\eta}}\left(\frac{c_i}{C}\right)^{-\frac{\eta-1}{\eta}}\right]\right] \tag{A.9}$$

where $\check{r} = \rho - \dot{\mu}_C/\mu_C$ and which gives (27). Next, I take the time derivative of (26):

$$\dot{\check{\mu}}_{h_i} = -\frac{1}{\xi}\left[\dot{p}_i - \delta_i\dot{\check{\mu}}_L\right], \tag{A.10}$$

and substitute this in (A.9) which gives

$$\check{\mu}_{h_i} = \frac{1}{\check{r}+\xi}\left[p_i\left[\frac{c_i}{h_i}\left[\frac{\theta}{\eta-1}-\left[\gamma+\frac{\theta}{\eta-1}\right]\left(\frac{h_i}{H}\right)^{\frac{\eta-1}{\eta}}\left(\frac{c_i}{C}\right)^{-\frac{\eta-1}{\eta}}\right]-\frac{1}{\xi}\frac{\dot{p}_i}{p_i}\right]+\frac{1}{\xi}\delta_i\dot{\check{\mu}}_L\right]. \quad \text{(A.11)}$$

I then substitute this result in (11) to find the following solution for the optimal price:

$$p_i = \delta_i\check{\mu}_L\left[\frac{1-\frac{\xi}{\check{r}+\xi}\frac{\dot{\check{\mu}}_L}{\check{\mu}_L}}{1+\frac{\xi}{\check{r}+\xi}\frac{c_i}{h_i}\left[\frac{\theta}{\eta-1}-\left[\gamma+\frac{\theta}{\eta-1}\right]\left(\frac{h_i}{H}\right)^{\frac{\eta-1}{\eta}}\left(\frac{c_i}{C}\right)^{-\frac{\eta-1}{\eta}}\right]-\frac{\xi}{\check{r}+\xi}\frac{\dot{p}_i}{p_i}}\right], \quad \text{(A.12)}$$

which can in turn be loglinearized to find (30).

A.2 Proofs

A.2.1 Proof to Lemma 2

First I take the time derivative of the loglinearized consumer demand function (18):

$$\dot{\tilde{c}}_i^R = -\eta\dot{\tilde{p}}_i^R + \theta\dot{\tilde{h}}_i^R. \quad \text{(A.13)}$$

Next, loglinearizing (6) allows me to write

$$\dot{\tilde{h}}_i^R = \xi\left[\tilde{c}_i^R - \tilde{h}_i^R\right]. \quad \text{(A.14)}$$

Then using (18), (A.13) and (A.14) in (25) and observing that under constant taxes, $\tilde{\tau}_i^R = \dot{\tilde{\tau}}_i^R = 0$, I find the following formula for change in $\tilde{c}_i^R$ as a function of $\tilde{c}_i^R$ and h_i^R:

$$\dot{\tilde{c}}_i^R = (\rho+\xi)\,\tilde{c}_i^R - \frac{\theta}{\eta-1}\left[(\rho+\xi)(\eta-1)+\xi(\theta-1)\right]\tilde{h}_i^R. \quad \text{(A.15)}$$

This gives the following system of dynamic equations:

$$\begin{bmatrix}\dot{\tilde{h}}_i^R\\ \dot{\tilde{c}}_i^R\end{bmatrix} = \begin{bmatrix}-\xi & \xi\\ -\frac{\theta}{\eta-1}\left[(\rho+\xi)(\eta-1)+\xi(\theta-1)\right] & \rho+\xi\end{bmatrix}\begin{bmatrix}\tilde{h}_i^R\\ \tilde{c}_i^R\end{bmatrix}. \quad \text{(A.16)}$$

From (21) and (23), $\dot{\tilde{h}}_i^R = -\xi\lambda\tilde{h}_i^R$ and $\dot{\tilde{c}}_i^R = -\xi\lambda\tilde{c}_i^R$, so

$$0 = \begin{bmatrix} -\xi(1-\lambda) & \xi \\ -\frac{\theta}{\eta-1}\left[(\rho+\xi)(\eta-1)+\xi(\theta-1)\right] & \rho+\xi(1+\lambda) \end{bmatrix} \begin{bmatrix} \tilde{h}_i^R \\ \tilde{c}_i^R \end{bmatrix}. \tag{A.17}$$

This gives

$$R^{mc}(\lambda) = (\theta-1)\left[(\rho+\xi)+\xi\frac{\theta}{\eta-1}\right]+\lambda\left[\rho+\xi\lambda\right], \tag{A.18}$$

with λ^{mc} implicitly determined by $R^{mc}(\lambda^{mc}) = 0$.

As $\theta < 1$, I know $R^{mc}(0) < 0$. Then as $\partial R^{mc}/\partial\lambda > 0$ for $\lambda > 0$ know there exists a solution $\lambda^{opt} > 0$. Next $R^{mc}(\lambda^{pc}) = -\lambda^{pc}\xi\theta\eta(\eta-1)^{-1}$, which is strictly negative for $\theta > 0$ and equal to zero if $\theta = 0$. Hence, I can conclude that if $\theta = 0$, $\lambda^{mc} = \lambda^{pc} = 1$, while if $\theta > 0$, $\lambda^{mc} > \lambda^{pc}$. Then for $\lambda' = 1+\frac{\theta}{\eta-1}$, $R^{mc}(\lambda') = \theta\frac{\eta}{\eta-1}\left[(\rho+\xi)+\xi\frac{\theta}{\eta-1}\right] > 0$, so $\lambda^{mc} < 1+\frac{\theta}{\eta-1}$ whenever $\theta > 0$. □

A.2.2 Proof to Lemma 3

The effect of a change in the individual parameters on λ^{mc} can be determined by taking a total differential of (A.18) and evaluating it at $\lambda = \lambda^{mc}$. I first consider the case with $\theta > 0$. Then $\frac{\partial\lambda^{mc}}{\partial\rho} = -\left.\frac{\partial R^{mc}/\partial\rho}{\partial R^{mc}/\partial\lambda}\right|_{\lambda=\lambda^{mc}}$, with $\partial R^{mc}/\partial\rho = (\theta-1)+\lambda$. As has been established in Section A.2.1, $\partial R^{mc}/\partial\lambda > 0$ and $\lambda^{mc} > 1-\theta$, so $\partial\lambda^{mc}/\partial\rho < 0$. Next, $\frac{\partial\lambda^{mc}}{\partial\xi} = -\left.\frac{\partial R^{mc}/\partial\xi}{\partial R^{mc}/\partial\lambda}\right|_{\lambda=\lambda^{mc}}$, where $\partial R^{mc}/\partial\xi = (\theta-1)\left(1+\frac{\theta}{\eta-1}\right)+\lambda^2$ which is ambiguous at first sight. Note however that by using (A.18) one can rewrite $\partial R^{mc}/\partial\xi = \frac{1}{\xi}\left(R^{mc}-\rho\left(\partial R^{mc}/\partial\rho\right)\right)$. At $\lambda = \lambda^{mc}$, $R^{mc} = 0$ and $\partial R^{mc}/\partial\rho > 0$, so $\partial R^{mc}/\partial\xi > 0$ and $\partial\lambda^{mc}/\partial\xi < 0$. Next, $\frac{\partial\lambda^{mc}}{\partial\eta} = -\left.\frac{\partial R^{mc}/\partial\eta}{\partial R^{mc}/\partial\lambda}\right|_{\lambda=\lambda^{mc}}$, with $\partial R^{mc}/\partial\eta = -(\theta-1)\xi\frac{\theta}{(\eta-1)^2} > 0$, which gives $\partial\lambda^{mc}/\partial\eta < 0$. In addition, $\frac{\partial\lambda^{mc}}{\partial\theta} = -\left.\frac{\partial R^{mc}/\partial\theta}{\partial R^{mc}/\partial\lambda}\right|_{\lambda=\lambda^{mc}}$, with $\partial R^{mc}/\partial\theta = \rho+\frac{\xi}{\eta-1}\left[\eta+2(\theta-1)\right]$. Rewriting this equation reveals that $\partial R^{mc}/\partial\theta < 0$, and thus $\partial\lambda^{mc}/\partial\theta > 0$, iff $(\eta-1) < (1-2\theta)\xi/(\rho+\xi)$. Inspection of $\partial R^{mc}/\partial\theta$ reveals this latter result continues to apply if $\theta = 0$.

Lastly, $\lambda^{mc} > 1$ if $R^{mc}(1) = (\theta-1)\left[(\rho+\xi)+\xi\frac{\theta}{\eta-1}\right]+[\rho+\xi] < 0$. This inequality can be rewritten as $\theta(\eta-1) < \theta(1-\theta)\xi/(\rho+\xi)$. □

A.2.3 Proof to Lemma 4

The optimal path of consumption can be found by combining (26), with (18), (A.13) and (A.14):

$$\dot{\tilde{c}}_i^R = (\rho+\xi)\tilde{c}^R - \left[\theta(\rho+\xi)+\xi\gamma(1-\theta)\right]\tilde{h}_i^R. \tag{A.19}$$

Then together with the time derivative of (18) I find the following system of dynamic equations

$$\begin{bmatrix} \dot{\tilde{h}}_i^R \\ \dot{\tilde{c}}_i^R \end{bmatrix} = \begin{bmatrix} -\xi & \xi \\ -[\theta(\rho+\xi)+\xi\gamma(1-\theta)] & \rho+\xi \end{bmatrix} \begin{bmatrix} \tilde{h}_i^R \\ \tilde{c}_i^R \end{bmatrix}. \tag{A.20}$$

From (21) and (23) I have $\dot{\tilde{h}}_i^R = -\xi\lambda\tilde{h}_i^R$ and $\dot{\tilde{c}}_i^R = -\xi\lambda\tilde{c}_i^R$, so

$$0 = \begin{bmatrix} -\xi(1-\lambda) & \xi \\ -[\theta(\rho+\xi)+\xi\gamma(1-\theta)] & \rho+\xi(1+\lambda) \end{bmatrix} \begin{bmatrix} \tilde{h}_i^R \\ \tilde{c}_i^R \end{bmatrix}. \tag{A.21}$$

This gives

$$R^{opt}(\lambda) = (\theta-1)[\rho+\xi(1-\gamma)] + \lambda(\rho+\xi\lambda), \tag{A.22}$$

with λ^{opt} implicitly determined by $R^{opt}(\lambda^{opt}) = 0$.

As $\theta < 1$, I know $R^{opt}(0) < 0$. Then as $\partial R^{opt}/\partial\lambda > 0$ for $\lambda > 0$, I know there exists a solution $\lambda^{opt} > 0$. Next, I can show that $R^{opt}(1-\gamma) = (\theta-\gamma)[\rho+\xi(1-\gamma)]$ while $R^{opt}(\lambda^{pc}) = (\theta-\gamma)(\theta-1)\xi$. Then if $\gamma = \theta$, we have $\lambda^{opt} = 1-\theta = \lambda^{pc}$. If $\gamma < \theta$, we must have $\lambda^{opt} \in (1-\theta, 1-\gamma)$ while if $\gamma > \theta$, $\lambda^{opt} \in (1-\gamma, 1-\theta)$. Finally, $R^{opt}(\lambda^{mc}) = -(\theta-1)\xi\left[\gamma + \frac{\theta}{\eta-1}\right]$ which is strictly positive if γ and/or θ are strictly positive, and equal to zero if $\gamma = \theta = 0$. Hence, I can conclude that $\lambda^{opt} < \lambda^{mc}$ if $\max\{\gamma,\theta\} > 0$, while $\lambda^{opt} = \lambda^{mc} = 1$ if $\gamma = \theta = 0$. □

A.2.4 Proof to Lemma 5

The effect of a change in the individual parameters on λ^{opt} can be determined by taking a total differential of (A.22) and evaluating it at $\lambda = \lambda^{opt}$. Then $\frac{\partial\lambda^{opt}}{\partial\gamma} = -\frac{\partial R^{opt}/\partial\gamma}{\partial R^{opt}/\partial\lambda}\Big|_{\lambda=\lambda^{opt}}$, with $\partial R^{opt}/\partial\gamma = \xi(1-\theta) > 0$. As has been established in Section A.2.3, $\partial R^{opt}/\partial\lambda > 0$, so $\partial\lambda^{opt}/\partial\gamma < 0$. Next $\frac{\partial\lambda^{opt}}{\partial\theta} = -\frac{\partial R^{opt}/\partial\theta}{\partial R^{opt}/\partial\lambda}\Big|_{\lambda=\lambda^{opt}}$, where $\partial R^{opt}/\partial\theta = \rho+\xi(1-\gamma) > 0$ so $\partial\lambda^{opt}/\partial\theta < 0$. Similarly, $\frac{\partial\lambda^{opt}}{\partial\rho} = -\frac{\partial R^{opt}/\partial\rho}{\partial R^{opt}/\partial\lambda}\Big|_{\lambda=\lambda^{opt}}$, with $\partial R^{opt}/\partial\rho = \lambda-(1-\theta)$. Then if $\gamma > \theta$, $\lambda^{opt} < 1-\theta$ and $\partial\lambda^{opt}/\partial\rho > 0$, if $\gamma < \theta$, $\lambda^{opt} > 1-\theta$ and $\partial\lambda^{opt}/\partial\rho < 0$, and if $\gamma = \theta$, $\lambda^{opt} = 1-\theta$, which gives $\partial\lambda^{opt}/\partial\rho = 0$. I know $\frac{\partial\lambda^{opt}}{\partial\xi} = -\frac{\partial R^{opt}/\partial\xi}{\partial R^{opt}/\partial\lambda}\Big|_{\lambda=\lambda^{opt}}$, where $\partial R^{opt}/\partial\xi = \lambda^2-(1-\theta)(1-\gamma)$. By (A.22), I can rewrite the latter as $\partial R^{opt}/\partial\xi = (R^{opt}-\rho(\partial R^{opt}/\partial\rho))/\xi$. At $\lambda = \lambda^{opt}$, R^{opt} is equal to zero, which implies $\partial\lambda^{opt}/\partial\xi$ is of opposite sign as $\partial\lambda^{opt}/\partial\rho$. Finally, as η does not appear in (A.22), $d\lambda^{opt}/d\eta = 0$. □

A.2.5 Proof to Proposition 2

First, the optimal adjustment to the steady state is equal to $\xi\lambda^{opt}$, while $\xi\lambda^{pc}$ is the adjustment rate to the steady state under constant taxes (including laissez-faire). From Lemma 4, it then directly follows that if $\gamma > (<)\theta$, $\xi\lambda^{opt} < (>)\xi\lambda^{pc}$. Next, define $\tilde{\tau}_i^{R,\lambda}(t)$ as the $\tilde{\tau}_i^R(t)$ that implements a given λ. Then, under perfect competition, by (22) and (24), $\tilde{\tau}_i^{R,\lambda}(t) = [(\theta - 1 + \lambda)/\eta]\,\tilde{h}_i^R(t)$ with $\tilde{h}_i^R(t) = \tilde{h}_i^R(0)e^{-\xi\lambda t}$. The term within square brackets is increasing in λ, and, by definition, equal to zero for $\lambda = \lambda^{pc}$. Thus, for $\lambda = \lambda^{opt}$, $\tilde{\tau}_i^{R,\lambda}(t)$ is positive (negative) and falling over time whenever $\lambda^{opt} > \lambda^{pc}$ and $\tilde{h}_i^R(0) > (<)0$. By Lemma 4 $\lambda^{opt} > \lambda^{pc}$ whenever $\gamma < \theta$. Similarly, for $\lambda = \lambda^{opt}$, $\tilde{\tau}_i^{R,\lambda}(t) < (>)0$ and rising (falling) over time if $\gamma > \theta$ and $\tilde{h}_i^R(0) > (<)0$ while $\tilde{\tau}_i^{R,\lambda}(t) = 0$ for all t if $\gamma = \theta$. □

A.2.6 Proof to Proposition 3

First, the optimal adjustment to the steady state is equal to $\xi\lambda^{opt}$, while $\xi\lambda^{mc}$ is the adjustment rate to the steady state under constant taxes (including laissez-faire). From Lemma 4, it then directly follows that if $\max\{\gamma,\theta\} > 0$, $\xi\lambda^{opt} < \xi\lambda^{mc}$. Next, define $\tilde{\tau}_i^{R,\lambda}(t)$ as the $\tilde{\tau}_i^R(t)$ that implements a given λ. Then, under monopolists, by (22) and (25), $\tilde{\tau}_i^{R,\lambda}(t) = \frac{1}{\eta}\left[\frac{(\theta-1)\left[(\rho+\xi)+\xi\frac{\theta}{\eta-1}\right]+\lambda(\rho+\xi\lambda)}{(\rho+\xi)+\xi\frac{\theta}{\eta-1}}\right]\frac{\rho+\xi}{\rho+\xi(1+\lambda)}\tilde{h}_i^R(t)$, with $\tilde{h}_i^R(t) = \tilde{h}_i^R(0)e^{-\xi\lambda t}$. The term within square brackets is increasing in λ, and, by definition, equal to zero for $\lambda = \lambda^{mc}$. By Lemma 4, if $\max\{\gamma,\theta\} > (=)0$, $\lambda^{opt} < (=)\lambda^{mc}$ and the bracketed term is negative (zero) for $\lambda = \lambda^{opt}$. As $t \to \infty$, $\tilde{h}_i^R(t)$ converges to zero, and so will $\tilde{\tau}_i^{R,\lambda}(t)$ for $\lambda = \lambda^{opt}$. □

A.3 Additional results

A.3.1 Closed-form solutions for λ^{mc} and λ^{opt}

From (A.18) and $R^{mc}(\lambda^{mc}) = 0$ I find the following closed-form solution for λ^{mc}:

$$\lambda^{mc} = \frac{\sqrt{\rho^2 - 4\xi(\theta-1)\left[(\rho+\xi)+\xi\frac{\theta}{\eta-1}\right]} - \rho}{2\xi}. \tag{A.23}$$

Similarly I find the following closed-form solution for λ^{opt} using (A.22) and $R^{opt}(\lambda^{opt}) = 0$:

$$\lambda^{opt} = \frac{\sqrt{\rho^2 - 4\xi(\theta-1)[\rho+\xi(1-\gamma)]} - \rho}{2\xi}. \tag{A.24}$$

A.3.2 Proof to $\partial(\xi\lambda^{opt})/\partial\xi > 0$

The effect of a change in ξ on the adjustment speed $\xi\lambda^{opt}$ is equal to $\partial(\xi\lambda^{opt})/\partial\xi = \lambda^{opt} + \xi(\partial\lambda^{opt}/\partial\xi)$. From Appendix (A.2.3) I know $\frac{\partial\lambda^{opt}}{\partial\xi} = -\frac{(\lambda^{opt})^2-(1-\theta)(1-\gamma)}{\rho+2\xi\lambda^{opt}}$. This allows me to write $\partial(\xi\lambda^{opt})/\partial\xi$ as $\frac{\partial(\xi\lambda^{opt})}{\partial\xi} = \frac{\lambda^{opt}[\rho+\xi\lambda^{opt}]+\xi(1-\theta)(1-\gamma)}{\rho+2\xi\lambda^{opt}}$, and conclude $\partial(\xi\lambda^{opt})/\partial\xi > 0$.

B Extension: habit formation and production externalities

In this appendix I show how the model can be adapted to explicitly account for the presence of an externality due to the production of one or multiple goods.[54] Optimally correcting the externality will require introducing a Pigovian tax (or subsidy), which in turn induces consumption to shift away from highly-taxed goods. I can then show that the optimal policy can be decomposed into two types of taxes; a time-independent externality tax τ^E that corrects the production externality, and a 'habit' tax τ_i^H that satisfies Propositions 1-3 and, as before, optimally manages the speed at which consumption moves toward the new steady-state.

Suppose that the production of a good i has an external effect on overall labor productivity, such that (7) is replaced by

$$L\left(1-\int_0^1 \Delta_i c_i di\right)=\int_0^1 \delta_i c_i di. \tag{B.1}$$

Here, Δ_i denotes the size of the external effect due to good i. If Δ_i is positive, the production of good i imposes a negative externality; the externality is positive if $\Delta_i < 0$ and absent if $\Delta_i = 0$. To ensure consumption remains bounded, I require $\delta_i + \Delta_i L > 0$ for all i.

Suppose that initially, no policy is in place to correct the externality. Then, equations (8)-(17) continue to apply. Steady-state C reads

$$C^* = \left[\int_0^1 \left(c_i^{R*}\right)^{\frac{\eta-1+\theta}{\eta}} di\right]^{\frac{\eta-1+\theta}{\eta}} \left[\int_0^1 \left(\delta_i + \Delta_i L\right) c_i^{R*} di\right]^{-1} L. \tag{B.2}$$

The presence of a negative externality ($\Delta_i > 0$) reduces overall labor productivity, and thereby the C^* that can be attained. Conversely, positive externalities ($\Delta_i < 0$), increase steady-state C^*. Concerning the transition to the steady state, one can show that equations (18)-(23) remain unchanged, which implies Lemmas 1-3 still apply.

When the production of certain goods imposes an externality on aggregate labor productivity, consumption patterns may not be first-best, even in steady-state. To determine the first-best consumption allocation, I again maximize (1), now subject to (2)-(6) and (B.1). This gives the following rule for optimal prices:

$$p_i(t) = \check{\mu}_L(t)\left[\delta_i + \Delta_i L\right] - \xi \check{\mu}_{h_i}(t), \tag{B.3}$$

where $\check{\mu}_{h_i}$ is still specified as in the main text (equation (27)). This in turn allows me to solve for

[54] As I consider a closed economy, production is equal to consumption, and the production externality can be reinterpreted as a consumption externality. To avoid any confusion with the consumption/habit internality, I will discuss the results in the context of a production externality only.

the optimal steady-state good i price:

$$p_i^* = \check{\mu}_L^* \left[\delta_i + \Delta_i L\right] \frac{\rho + \xi}{\rho + \xi\,(1-\gamma)}. \tag{B.4}$$

In can then prove the following proposition, which replaces Proposition 1:[55]

Proposition B.1. *In steady state, laissez-faire consumption choices are optimal if only if* $\Delta_i/\delta_i = \Delta_b/\delta_b$ *for all* $i,b \in [0,1]$. *A relative tax that satisfies* $\tau_i^{R*} = \frac{1+(\Delta_i/\delta_i)L}{1+(\Delta_b/\delta_b)L}$ *implements optimal steady-state consumption.*

Proof. By (B.4), the optimal relative price in steady state satisfies $p_i^{R*} = \frac{\delta_i + \Delta_i L}{\delta_b + \Delta_b L}$. This is equal to (17) iff $\tau_i^{R*} = \frac{1+(\Delta_i/\delta_i)L}{1+(\Delta_b/\delta_b)L}$. Under laissez faire, $\tau_i = 1$ for all i, so $\tau_i^{R*} = 1$. Then the laissez-faire τ_i^{R*} is equal to the first-best τ_i^{R*} iff $\Delta_i/\delta_i = \Delta_b/\delta_b$ for all $i,b \in [0,1]$. □

To determine the optimal path of consumption, prices and habits as consumption transitions away from good i, I solve for $\lambda^{E,opt}$, the λ for the optimal path in the presence of the production externality. This $\lambda^{E,opt}$ can then be used in (31) and (32) to determine the value of $\tilde{\tau}_i^R(t)$ required to implement the optimal transition path under perfect competition and monopolistic supply, respectively. I find that the optimal rate of adjustment is independent of the production externality:

Lemma B.1. $\lambda^{E,opt} = \lambda^{opt}$

Proof. Loglinearizing (B.3) around the steady state gives (30). (18)-(23) still apply, so $\lambda^{E,opt} = \lambda^{opt}$. □

The result that $\lambda^{E,opt} = \lambda^{opt}$ implies that Propositions 2 and 3 continue to apply, and the $\tilde{\tau}_i^R$ required to implement the optimal consumption path under perfect competition and monopolistic supply are still characterized by (31) and (32) with $\lambda = \lambda^{opt}$, respectively. Put differently, even though the production externality affects the tax that implements optimal steady-state consumption, it does not affect the optimal adjustment path to the steady state and thus the path of taxes, $\tilde{\tau}_i^R$, required to implement it. Hence, the optimal corrective tax can be separated into two parts. The first part corrects for the external effect on labor productivity. This tax is time-independent. The second part corrects for habit formation, and changes as the economy moves to its steady-state.

Starting in a steady state with no taxes, the introduction of an optimal corrective tax then sets in motion a process where consumption shifts away from goods with a relatively high Δ_i. In this context, the optimal $\tilde{\tau}_i^R(t)$ can be interpreted as the discount (or premium) to the optimal long-run tax that ensure the economy moves toward the new steady-state consumption bundle at its optimal rate. For example, suppose we start in a steady state with no taxes and $\frac{1+(\Delta_i/\delta_i)L}{1+(\Delta_b/\delta_b)L} > 1$, such that

[55] In fact, Proposition 1 can be considered a special case of Proposition B1.

the optimal steady-state relative tax is positive ($\tau_i^{R*} > 1$). Suppose also that for any $\tilde{h}_i^R(t) > 0$, the optimal $\tilde{\tau}_i^R(t)$ is negative. Then, the optimal tax policy would be to introduce a positive tax which rises over time to the long-run optimal rate τ_i^{R*}.

C Discrete-time setup

Instead of discretizing the pricing rules (11) and (26) (or (B.3)), I rederive them by solving the discrete-time model 'bottom up'. I solve the more general setup presented in Appendix B. By setting all $\Delta_i = 0$, all results can be directly compared to their counterparts in Sections 3 to 4. Note that in this setup, as discussed in Appendix B, all Lemmas, as well as Propositions 2 and 3, continue to apply and we can interpret the numerical results referring to these analytical results.

In discrete time, (2)-(5), (B.1) and (8)-(9) still apply. Welfare is replaced by

$$W_t = \sum_{v=t}^{v=\infty} (1+\rho)^{-(v-t)} U_v, \tag{C.1}$$

while the equation of motion for habits now reads

$$h_{it+1} = \xi c_{it} + (1-\xi) h_{it}, \tag{C.2}$$

where the use of time subscripts indicates we now deal with the discrete-time version of the model. Under perfect competition, $p_{it} = \delta_i p_{Lt} \tau_{it}$ still. The good i monopolist maximizes $\Pi_{it} = \sum_{v=t}^{v=\infty} \left[c_{iv} \left(p_{iv} - \delta_i p_{Lv} \tau_{iv} \right) \prod_{x=t+1}^{x=v} \frac{1}{1+r_x} \right]$, subject to (8) and (C.2) by choosing $[c_{iv}]_{v=t}^{v=\infty}$. This gives the following pricing rule:

$$p_{it} = \delta_i p_L \tau_{it} \frac{\eta}{\eta - 1} \left[\frac{1 - \frac{1-\xi}{r_{t+1}+\xi} \frac{p_{Lt+1}\tau_{it+1} - p_{Lt}\tau_{it}}{p_{Lt}\tau_{it}}}{1 + \frac{\xi}{r_{t+1}+\xi} \frac{\theta}{\eta-1} \frac{c_{it+1}}{h_{it+1}} \frac{p_{it+1}}{p_{it}} - \frac{1-\xi}{r_{t+1}+\xi} \frac{p_{it+1}-p_{it}}{p_{it}}} \right]. \tag{C.3}$$

The interest rate r_{t+1} is determined by the consumption Euler equation: $\frac{dU_t/dC_t}{\mathbb{E}^{rc}[dU_{t+1}/dC_{t+1}]} = \frac{1+r_{t+1}}{1+\rho}$. As pointed out in footnote 18, further assumptions must be made regarding the representative consumer's anticipation of the change in H as the economy transitions to the steady state. For the numerical application, I have computed the transition for two cases; one where the consumers anticipate the change in H, and one where they do not. As results are virtually indistinguishable I only present the latter.[56]

The policymaker instead chooses the $[c_{iv}]_{i\in[0,1],v=t}^{v=\infty}$ that maximize (C.1) subject to (2)-(5), (B.1)

[56] This is not surprising. The shock I consider is one that is relatively small, and applies to a small share of goods. Hence, even if shifts in c_i^R are large, shifts in C and H, and hence r, are expected to be small.

and (C.2). This gives the optimal price for good i

$$p_{it} = \check{\delta}_i \check{\mu}_{L_t} \left[\frac{1 - \frac{1\text{-}\xi}{\check{r}_{t+1}+\xi} \left(\frac{\check{\mu}_{L_{t+1}} - \check{\mu}_{L_t}}{\check{\mu}_{L_t}} \right)}{1 + \frac{\xi}{\check{r}_{t+1}+\xi} \frac{p_{it+1}}{p_{it}} \frac{c_{it+1}}{h_{it+1}} \left[\frac{\theta}{\eta-1} - \left[\gamma + \frac{\theta}{\eta-1} \right] \left(\frac{c_{it+1}}{C_{t+1}} \right)^{-\frac{\eta-1}{\eta}} \left(\frac{h_{it+1}}{H_{t+1}} \right)^{\frac{\eta-1}{\eta}} \right] - \frac{1\text{-}\xi}{\check{r}_{t+1}+\xi} \left(\frac{p_{it+1} - p_{it}}{p_{it}} \right)} \right], \tag{C.4}$$

where I define $\check{\delta}_i \equiv \delta_i + \Delta_i L$, and $\check{r}_{t+1}$ is now determined by the discrete-time Euler equation $\frac{\partial U_t / \partial C_t}{\partial U_{t+1} / \partial C_{t+1}} = \frac{1+\check{r}_{t+1}}{1+\rho}$. A comparison of (C.3) to (11), and (C.4) to (26), reveals two differences between the continuous and discrete-time pricing rules.

First, in continuous time, the denominator features the instantaneous ratio of good-specific consumption to habits, $c_i(t)/h_i(t)$, while in discrete time this is the $t+1$ ratio c_{it+1}/h_{it+1}, multiplied by the ratio of time $t+1$ to time t good-specific prices. Similarly, for the discrete-time optimal price (C.4), C and H are evaluated at time $t+1$, while in continuous time we have $C(t)$ and $H(t)$. This can be explained as follows. In continuous time, the habit adjustment occurs instantaneously. Hence to evaluate the value of the habit, instantaneous consumption, habits and prices are relevant. In discrete time, it is the next-period habit that adjusts, and thus next-period consumption, habit and prices are used to determine the value of investing in the habit. In both cases, the value of the habit is then evaluated relative to current prices.

Second, in the discrete-time pricing rules, the future change in taxes and prices are multiplied by an additional $1-\xi$. This is intuitive. In the discrete-time model, if $\xi = 1$, habits fully adjust from one period to another and the decision maker only needs to know the value of habits one period ahead. Hence, for $\xi = 1$, future changes in the value of the habit, captured by the change in prices (net of taxes), become irrelevant and drop out. In the continuous-time pricing rules, (11) and (26), this full adjustment from one *instant* to another occurs if $\xi \to \infty$. Here again, future price and tax changes drop out. Note that where any $\xi \geq 0$ can be rationalized in the continuous time model, in the discrete-time model only $\xi \in [0, 1]$ are sensical.

C.1 Additional details for numerical results

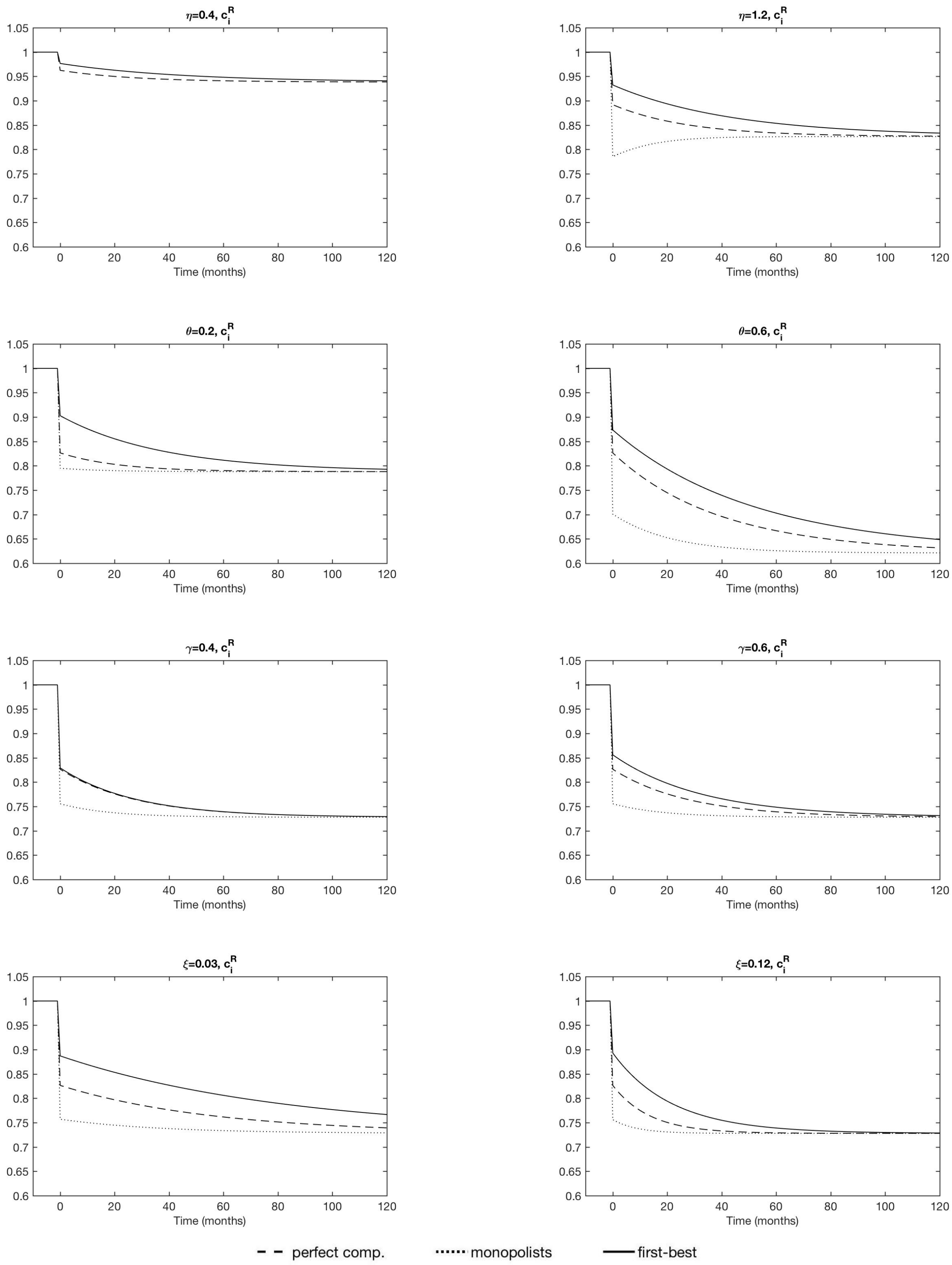
I use the Dynare package version 4.4 to solve the model numerically.

I assume the charge optimally corrects for a production externality, as in Appendix B. Put differently, I assume that for 'unhealthy foods', $\Delta_i = 0.1$ while $\Delta_i = 0$ for all other goods (see Proposition B1 and (C.4)), and that in the initial steady-state relative consumption and habits equal unity. The difference between the welfare gains including and excluding the benefit from correcting the production externality comes from a difference in initial steady states. In the former case, I assume $\Delta_i = 0.1$ already prior to the implementation of policy, implying the initial steady state was distorted. In the latter, I consider an initial steady state that was undistorted: prior to $t = 0$, $\Delta_i = 0$ for all i. Here, the charge is implemented immediately in response to an increase in Δ_i to 0.1 for $i \in [0,n]$ at $t = 0$. Initial relative prices and consumption, as well as the new steady state are identical in both cases, yet due to the noninternalized production externality, initial consumption levels are lower in the former case.

All transition paths presented in this paper resemble the case where the initial steady state is distorted. To determine gains for the alternative case, I compute all transition paths again. Results for c_i^R, p_i^R and τ_i^{opt} are virtually indistinguishable.

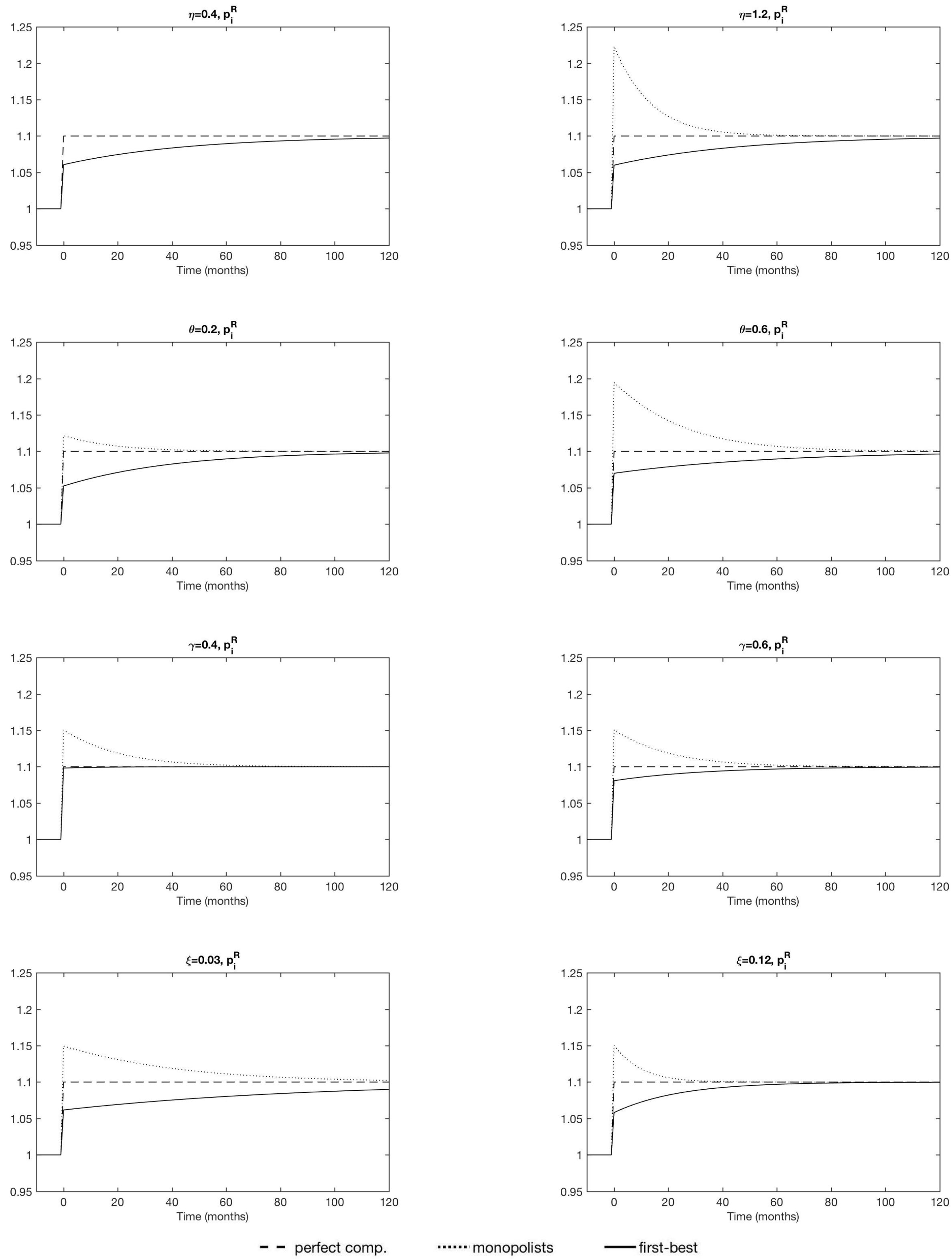
D Additional figures

Figure D.1: Sensitivity analysis, c_i^R



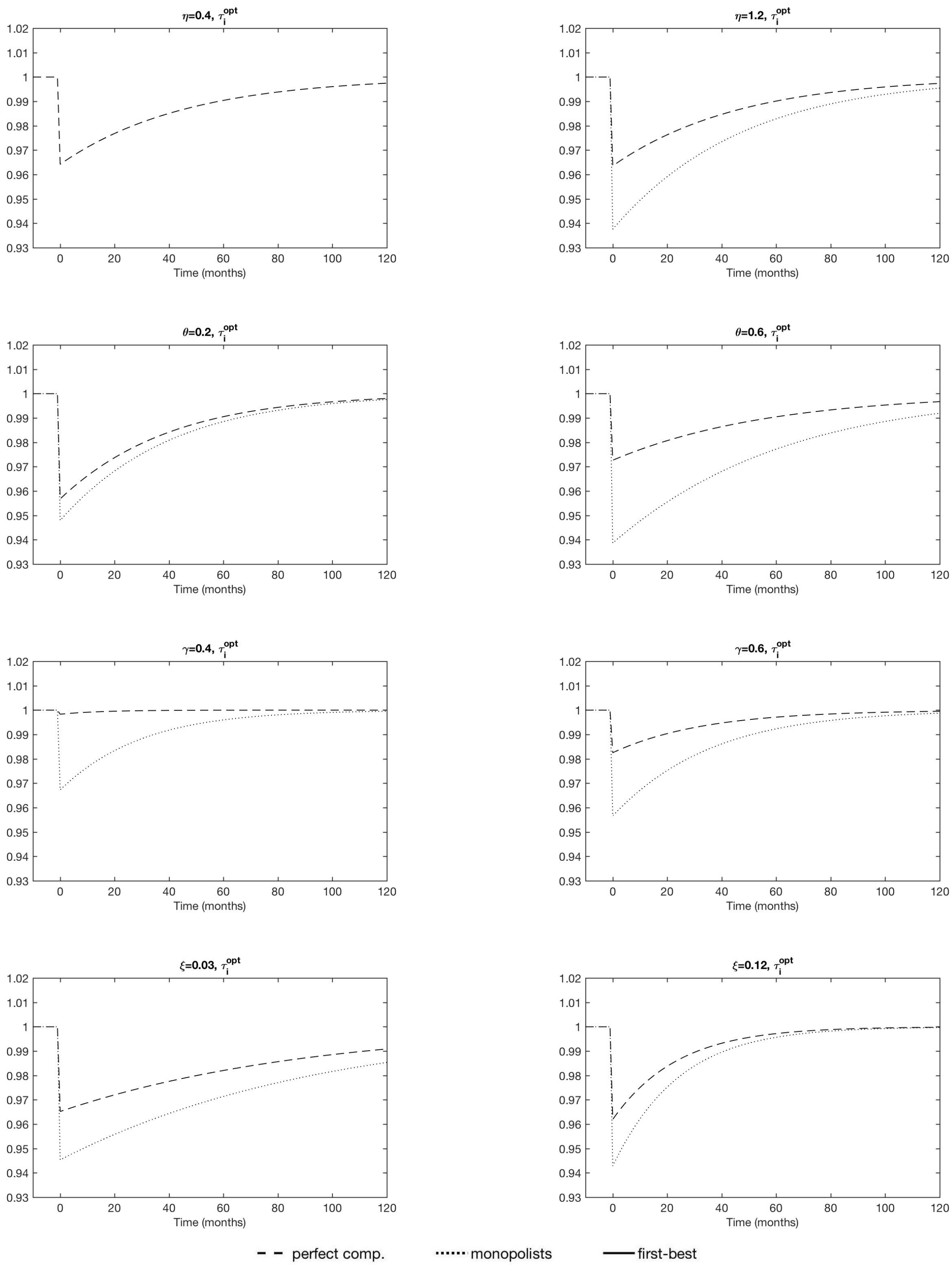
The curves depict responses to the unanticipated introduction of the 10 percent charge at $t = 0$, where I assume the economy is in steady state for all $t < 0$. Responses are shown for a good $i \in [0,n]$ relative to any good $b \in (n,1]$.

Figure D.2: Sensitivity analysis, p_i^R



The curves depict responses to the unanticipated introduction of the 10 percent charge at $t = 0$, where I assume the economy is in steady state for all $t < 0$. Responses are shown for a good $i \in [0, n]$ relative to any good $b \in (n, 1]$.

Figure D.3: Sensitivity analysis, τ_i^{opt}



The curves depict responses to the unanticipated introduction of the 10 percent charge at $t = 0$, where I assume the economy is in steady state for all $t < 0$. Responses are shown for a good $i \in [0, n]$ relative to any good $b \in (n, 1]$.