Rational Addiction to Soda by Obesity Status

Josh Reed, Yizao Liu, Emily Wang, Eliana Zeballos, Edward C. Jaenicke

Selected Paper prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

Rational Addiction to Soda by Obesity Status

Josh Reed, Penn State University

Yizao Liu, Penn State University

Emily Wang, University of Massachusetts Amherst

Eliana Zeballos, USDA Economic Research Service

Edward C. Jaenicke, Penn State University

Abstract: Assessing food addiction may play an important role in implementing policy to tackle obesity in the United States. Becker and Murphy's (1988) theory of rational addiction is commonly employed in the economics literature to assess such behavior. This paper assesses the use of both standard empirical models of rational addiction (Becker, Grossman, and Murphy, 1994) as well as recent empirical refinements (Dragone and Raggi, 2021). We model the consumption of regular soda which is of great interest in a food addiction context using a rich consumer panel provided by IRI (Information Resources Inc.). We show evidence that rational addiction models which take advantage of recent empirical refinements are effective in not identifying non-addictive products as addictive where the canonical models consistently show strong evidence of rational addiction to low-fat/skim milk. Using this empirical refinement, we do not show strong evidence of rational addiction to regular soda, although there is slightly stronger evidence for obese households. We present a novel random coefficient modelling approach to rational addiction which better accounts for consumer heterogeneity in addictive behavior and broadly confirms the results of fixed coefficient modelling approaches. This strategy finds that only 1.01% of households can be described as rationally addicted to soda.

This work is supported by USDA's National Institute of Food and Agriculture (NIFA). The findings and conclusions in this publication are those of the authors. The analysis, findings, and conclusions expressed in this report should not be attributed to IRI. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

I. Introduction

When rats were given a mutually exclusive choice between sweetened water and intravenous cocaine 94% consistently preferred the sweetened water (Lenoir et al. 2007). This and other studies have suggested a strong potential for addiction to common food products, especially those high in sugar like soda, in humans. The prevalence of obesity in the United States in 2017–2018 was 42.4 percent among adults aged 20 and over, and it has increased from 30.5 percent in 1999–2000, part of this increase could be a result of addictive consumption patterns (CDC, 2021; Alston and Okrent, 2017; Ogden et al., 2015). Relying on models of rational addiction (e.g., Becker and Murphy, 1988), recent studies look for evidence of addictive behavior among certain foods, and explore the associated policy implications (e.g., Richards, Patterson, and Tegene, 2007, Zhen et al., 2011; Gordon and Sun, 2015). There is evidence to suggest that policies aimed at rationally addicted consumers can be effective if they increase future costs associated with addiction, so consumers become rationally incentivized against feeding the addictive behavior today. With respect to food addiction and obesity, this lesson implies that policies which tax a particular food or beverage (such as soda), or a particular nutrient (such as sugar) might be more effective in reducing consumption than a static treatment of demand would imply.

This paper utilizes a rich panel of household level purchase data provided by IRI to look for evidence of rational addiction to regular soda and leverages the non-addictiveness of low-fat/skim milk to ensure that the modelling strategy correctly identifies non-addictive products. We first utilize the canonical, AR(2), model of rational addiction (e.g., Chaloupka et al, 1991; Becker, Grossman, and Murphy, 1994) and show that regular soda exhibits rational addiction. However, we also show that this same model predicts rational addiction to low-fat/skim milk which should be plainly non-addictive, this is a well-documented problem of these models (Auld and Grootendorst, 2004). This problem has recently been addressed by Dragone and Raggi (2021) who show that their updated AR(1) model of rational addiction solves this “milk addiction paradox”. Applying this model, we show that neither regular soda nor low-fat/skim milk show strong evidence of rational addiction. However, we see stronger, although still not statistically significant, evidence of rational addiction to regular soda in households where the household’s average adult BMI categorizes them as “obese”.

Taking advantage of our micro level data we propose a novel random-coefficients strategy for estimating both the canonical AR(2) and more recent AR(1) models of rational addiction. This estimation procedure more accurately represents addiction as a phenomenon which is heterogeneously distributed within a population. Additionally, this strategy allows us to individually identify which households show evidence of rational addiction. We once again confirm the viability of the AR(1) model in solving the milk addiction paradox because while the AR(2) model predicts unreasonably large (upwards of 20%) portions of the population are addicted to both regular soda and low-fat/skim milk, the updated AR(1) model predicts that only 1.01% of households show significant evidence of rational addiction to regular soda. There is also stronger evidence of rational addiction to soda in the obese BMI group consistently across all models.

Overall, this research implies that rational addiction is not an effective framework for assessing demand for regular soda. However, we present another model which we plan to estimate in future which could show evidence to the contrary. Gordon and Sun’s (2015) dynamic model of addiction and stockpiling could help to disaggregate the effects of addiction and stockpiling. Stockpiling has the opposite implications of past purchases on current consumption as addiction, if a household bought more in the past to stockpile they would buy less today. We intend to estimate this model in later revisions of this work.

II. Literature Review

Becker and Murphy's (1988) seminal work on rational addiction describes many behaviors commonly associated with addiction in a highly tractable model. The term rational addiction may seem at first an oxymoron, how indeed could the erratic and self-destructive behaviors of the alcoholic, heroin user, or the overeater be described as rational? However, Becker and Murphy's model describes forward-looking utility maximizing individuals with stable preferences that economists would describe as rational. The process of addiction is described where past consumption affects future consumption in a process of "learning by doing", addiction is thus inherently a dynamic process. A necessary condition for addiction is that past consumption of a good raises the marginal utility of present consumption; therefore, consumption increases with a reinforcing effect over time. Therefore, the model predicts a strong correlation between past consumption and future consumption known as adjacent complementarity or reinforcement. Where rational addiction differs from the older theory of myopic addiction is that not only past consumption can affect current utility, but future consumption can also have significant effects on current utility and thus demand. If a rational addict knows the good they are addicted to will become more expensive in the future due to increased taxes, prices, or prohibition they may be inclined to mediate their current consumption, so they don't need as much of it in the future. As the model predicts forward-looking, utility maximizing behavior it also predicts that current consumption will be sensitive to future consumption and/or prices. Future price expectations can have variable effects on future consumption. Indeed, the theory predicts both weak addictions and strong addictions. Weak addictions, those we should commonly expect to cigarettes or soda, can be mediated by expectations of a future price increase. Strong addictions, like those common to heroin addicts, are bolstered by expectations of future price, the logic being that one wants to enjoy something while they can still afford it, these addictions can only be ended by quitting cold turkey (Becker and Murphy, 1988). When empirically estimating rational addiction researchers are generally looking for A) significant positive effects of past consumption on current consumption and B) significant effects of future consumption and/or prices on current consumption.

If a good exhibits rational addiction then the Becker and Murphy model (1988) suggests an important policy point. For excise taxes to be most effective they should impose a permanent price change, the rational consumer, realizing that the costs of future consumption have risen dramatically will then substantially curtail their consumption. Sometimes this recommendation is inverted to suggest that policymakers should apply permanent subsidies to "beneficial" addictions, such as theatergoing, to encourage them (Concetta and Infante, 2016).

While proponents of the Becker and Murphy (1988) model of rational addiction praise its tractability and flexibility, its consistency with rational choice theory, and its predictive power there are certainly detractors. Rogeberg (2004) for instance points out the absurdity of the assumption of the rational forward-looking addict. The theory of rational addiction effectively assumes that people can predict the future with regards to a product they become addicted to, they know how much they will be addicted to it in the future and what harmful effects it will have. Additionally, we assume that rational addicts are perfected, deliberate, and conscious decision makers. Rational addiction certainly presents some powerful abstractions; however, many researchers have found models of rational addiction useful and empirical refinements have made their results more believable.

Early empirical specifications of rational addiction focused on addiction to cigarettes (e.g., Chaloupka, 1991; Becker, Grossman, and Murphy, 1994). These models derive empirical demand estimations from first order conditions of utility functions presented in Becker and Murphy (1988). These derivations are discussed in more detail in the methodology section of this paper. These early works confirmed some predictions of the rational addiction model, namely that smoking is addictive, that addiction to cigarettes is rational, not myopic, in that consumers are forward looking, and that more educated and older individuals' addiction behavior is more rational (Chaloupka, 1991; Becker, Grossman, and Murphy, 1994).

Since the publication of this earlier work confirmed many of the hypotheses of the rational addiction model many researchers sought to apply it to a broader range of products. The demand for cocaine has been shown to exhibit rational addiction with cocaine consumers adjusting their current consumption down in anticipation of future price increases resulting from DEA crackdowns (Grossman and Chaloupka, 1998). However, the framework of rational addiction does not seem to apply to all addictive drugs in all settings, Liu et al (1999) find that in the case of the opium market in Taiwan 1914-1942 rational addiction is not an effective framework for describing consumption decisions where a myopic addiction model, one where only past and not future consumption does not significantly predict current consumption, is a more appropriate framework. Rational addiction models have since been extended to examine addiction to coffee, liquor, theatre attendance, driving, and many other potentially addictive commodities and activities (Olekalns and Bardsley, 1996; Baltagi and Griffin, 2002; Concetta and Infante, 2016; Collet, Lapparent, and Hivert, 2015 respectively).

However, later research has also shown that these same models could be used to show rational addiction in many non-addictive product categories, like dairy milk (Auld and Grootendorst, 2004). Recently this issue has been addressed, Dragone and Raggi (2021) show that the standard rational addiction model based off first-order conditions of utility functions in Becker and Murphy (1988) leads to explosive consumption paths while their updated model follows a stable saddle path that smoothly converges to a steady state level of consumption. The derivation of this model is also discussed in more detail in the methodology section. Using this model Dragone and Raggi (2021) show evidence of addictive demand for cigarettes but not for non-addictive products like oranges, eggs, and milk. Taking insights from both Becker, Grossman, and Murphy's (1994) and Dragone and Raggi's (2021) models we estimate a model of addictive demand which we apply to various classes of food products.

A related issue is the estimation of the implied rate of time discount, often standard models of rational addiction will produce unrealistic estimates which vary wildly in the literature (Laporte, Dass, and Ferguson, 2017). Laporte, Dass, and Ferguson (2017) find that the strange values of the implied discount rate often found in the literature may result from saddle-point dynamics associated with individual level inter-temporal optimization problems. They briefly refer to the related discussion of quasi-hyperbolic discounting versus exponential discounting and the assumption of time-consistent preferences in the rational addiction model (Laporte, Dass, and Ferguson, 2017). This issue is expanded upon by Piccoli and Tiezzi (2021) who embed quasi-hyperbolic discounting into the standard rational addiction model. This allows them to derive a test to distinguish between time-consistent versus time-inconsistent naïve agents. They do not however reject the hypothesis that Russian smokers are time-consistent in their demand for cigarettes (Piccoli and Tiezzi, 2021).

While the standard rational addiction framework is fairly simple to estimate, often requiring little more than OLS or 2SLS, more structural models using random coefficients mixed logit models or a dynamic model of rational addiction with endogenous consumption and stockpiling have been developed (Richards Paterson and Tegene, 2007; Gordon and Sun, 2015).

Of special interest to this project is Richards Patterson and Tegene (2007) who use a dynamic coefficients mixed-logit model (as in Berry Levinsohn and Pakes, 1995) to describe rational addiction to fat, protein, carbohydrates, and sodium using Nielsen homescanner data. This specification allows the analysis of substitution patterns between nutrients which is not possible at the product level due to the sheer number of products that consumers regularly purchase. While their estimation differs from standard models of rational addiction their implications remain the same. First, habit persistence describes the propensity of consumers to consume more of products they consumed in the past, as in a myopic addiction model as well as a rational addiction model. More importantly for the case of rational addiction, they also measure where a reduction in future consumption can reduce utility and thus consumption in the current period. Richards, Patterson and Tegene (2007) find that all the nutrients they examine show evidence of rational addiction, especially carbohydrates. However, their sample is very small only including 30 households and does not examine other factors related to nutrient addiction such as obesity.

Additionally, Gordon and Sun (2015) present a dynamic model of addiction with endogenous consumption and stockpiling. This treatment closely follows the approach of Hendel and Nevo (2006a) to identify stockpiling behavior. Addiction in this model is treated similarly to a physical stock of goods, this "stock of addiction" represents how past consumption affects present marginal utility and decays over time. To maintain conceptual consistency with Becker and Murphy (1988), Gordon and Sun (2015) model consumers expectations of future prices and explicitly model how future prices affect present utility in both the processes of addiction and stockpiling. This model considers three goods cigarettes, crackers, and butter and considers attempts to fit models containing addiction, stockpiling, and both dynamic demand shifters together. While consumption of butter and crackers is well modeled by the models which consider only stockpiling the consumption of cigarettes is best modeled by the model containing both addiction and stockpiling. This model is powerful because it can differentiate between addiction and stockpiling, two dynamic processes which both can have intertemporal effects on consumption. As such we intend to implement it in later iterations of this work, and it is described in detail in the methodology section of this paper.

The literature concerning rational addiction is broad and varied. Many recently published empirical refinements (e.g., Dragone and Raggi, 2021;Picoli and Tiezzi, 2021; Laporte, Dass, and Ferguson, 2017; Gordon and Sun, 2016) show that Becker and Murphy's (1988) theory of rational addiction is a topic of great interest to researchers still today.

The primary goal of this project is to estimate a model of addiction to apply in the context of a large panel of consumer purchase decisions provided by IRI. For this study we estimate the standard empirical model of rational addiction specified by Becker, Grossman, and Murphy (1994) as well as applying the empirical refinement suggested by Dragone and Raggi (2021) to both regular soda and low-fat/skim milk. We present a novel random-coefficients strategy which takes advantage of the micro scale of our data. Additionally, we lay the groundwork for estimation of a dynamic model of rational addiction and stockpiling as in Gordon and Sun (2015).

III. Data

Data from this project were collected by Information Resources Inc. (IRI) and provided through the United States' Department of Agriculture, Economic Research Service (details on data can be found in ERS, 2021). Data include household level data and store level data on food purchases. This dataset is comparable to the Nielsen scanner dataset. This project focuses on data collected between 2013 and 2018, although data from between 2008 and 2012 is also available. While we have found data from 2019 and 2020 to be somewhat incomplete at this point we do plan to use it in the future as well as any future data that comes out during this project. This project so far focuses on two specific products regular soda which is of particular interest in a food addiction context and low-fat/skim milk which should be blatantly non-addictive and thus provides a point of comparison. Although, we have also seen some interesting results for products such as diet soda, whole milk, ice-cream, potato chips, and tortilla chips which we intend to explore in more detail in the future. Because the dataset remains quite large we can use a consistent panel including only households which: A) are in the sample for at least 5 years, B) are members of IRI's static panel, C) consistently consume the product in question at least 4 times a year, and D) participate in the med-profiler survey.

Household level data includes both the IRI Consumer Network household scanner data and supplemental Medprofiler data, these data were compiled for the entire U.S.. The Consumer Network household scanner data includes information on purchase quantities, spending, and product attributes for each grocery trip a consumer makes throughout their participation in the survey. We aggregate this data by household and product categories monthly to estimate monthly purchases and spending, purchase units are listed in fluid ounces. We restrict this sample to the "static panel" a subset of consumers identified by IRI who consistently record their purchases, this cuts the sample by about half. Table 1 shows the average number of households in our sample is 5,437 for regular soda and 9,134 for low-fat/skim milk. The primary reason for this difference in samples is that we limit the sample to households which consume at least four times a year for estimation purposes.

To account for variation in consumption due to household size and the ages of various household members we measure consumption per "adult equivalent" household member (World Bank, 2005). This measure effectively downweighs children in the calculation of family size to account for their lesser caloric needs. In this study we count children ages 0-6 as 0.2 adult equivalents, those ages 7-12 as 0.3, and those ages 13-17 as 0.5. As shown in Table 1 the average proportion of households with children is around 26% for our regular soda sample and 20% for our low-fat/skim milk sample which corresponds to the average household size being a little higher than two. As our sample is restricted to households which consistently consume the product it is interesting that the soda sample contains a larger portion of households with children.

To accurately measure consumers consumption decisions over time we must also note when consumers do not purchase a product. After first finding the set of consumers who purchase a given product at least once over the sample period to comprise our sample group we then zero-fill the data on consumption in the months in which we do not observe consumption of that product. We find that for estimation purposes it is best to include only households which consume the product in question at least four times a year every year.

Accounting for price has been especially problematic in this environment, where we also want to know the price available to consumers when they do not make a purchase, as well as accounting for heterogeneity in consumers tastes, store choices, and timing. Thus, we construct a measure of price which is specific to each consumer in that it is based on A) the chains they are observed to visit. B) the region of the country they live in, C) the brand and size they typically purchase, and D) when they make the purchase. To begin constructing this measure we list every store visit IRI participants log in the data, this along with its date and the chain that the participant visits. We then combine this information with every purchase of the relevant product, be it low-fat/skim milk or regular soda, along with the brand and size chosen the quantity purchased and the amount paid. From here we can identify some key pieces of information, first for each consumer in each year we identify the modal brand and modal size of their purchase decisions. We also identify the average price for each brand size combination at each chain in each region for each week. Matching average prices for each brand, chain, region, and week to participants observed grocery trips and modal brands and sizes thus accurately describes the price environment that consumers face, even when they do not make a purchase. However, this method has been somewhat incomplete due to the level of specificity when selecting on specific brands. Thus, we also aggregate brands into "tiers", for regular soda these tiers are based on the parent company of the specific soda brand[1] (Coca-Cola, PepsiCo, Dr Pepper Snapple Group/Keurig Dr Pepper, private label, or other) for low-fat/skim milk tiers are based on whether the milk is organic or not. When we do not observe the average price for a specific brand we assume the consumer observes the average price of their modal brands "brand tier". Overall, this measure of price is highly correlated with the observed prices that consumers pay when they do make a purchase. For estimation we find the average price per fluid ounce per month for each household, with 98.0% of household-months for regular soda and 99.8% of household-months for low-fat/skim milk having a price identified under this definition. While this measure of price is complex to estimate it provides an accurate view of the prices that participants observe, even when they do not make a purchase, which accounts for heterogeneity in consumers tastes, store choices, and timing.

At the consumer level we also have access to the Medprofiler survey which surveyed a subsample of consumers annually in the broader Consumer Network survey. This data crucially includes data for all household members including not only their height and weight but also qualitative descriptions of a variety of physical and mental health conditions including diabetes, heart disease, anxiety, and depression. This project so far uses only the information for height and weight to calculate body mass index (BMI). Aggregating BMI to the household level is quite problematic as it is difficult to assess an individual's contribution to the household's collective health. Thus, we chose to use average adult BMI in the household as the least problematic and simplest to implement method of aggregation. BMI categories are assigned consistently with CDC's methods where average household adult BMI<25 is assigned to the "normal" category (our analysis is mostly focused on obesity, so we include households in the "underweight" category in the "normal" category), average household BMI>30 is assigned to the "obese"

[1] This methodology requires researching many individual brands and is extremely time-consuming, thus constituting the primary roadblock for estimation on a wider range of products. We categorize the parent companies of brands representing the top 99% of regular soda sales.

category and anything in-between is assigned to the "overweight" category (CDC, 2020). As Table 1 shows the average BMI of adults in our sample is around 29 which fits within the category of "overweight".

IV. Methods

The primary goal of this project is to find the best model of addiction and apply it to food and beverage products, we test the validity of each model in the context of the IRI data with repeated purchase decisions by individual consumers. Given the prominence of rational addiction models in the economic literature we first investigate a series of rational addiction models. In specific, we investigate the model in Becker, Grossman, and Murphy (1994) which forms the basis of much of the literature concerning rational addiction, we call this class of models AR(2) because consumption in two different time periods, lag and lead, are included as covariates. We also investigate a recent finding by Dragone and Raggi (2021) which proports to solve the problem of AR(2) models routinely identifying non-addictive products as addictive, such as milk (Auld and Grootendorst, 2004). This model includes only past consumption as a covariate and assesses consumers forward looking behavior by including future price as a covariate, so we call this model AR(1).

We select products for this sample to rule out models which show addiction to plainly non-addictive products. To test this proposition, we include low-fat/skim milk, which has been shown to display rational addiction using the AR(2) specification but not the AR(1) specification (Auld and Grootendorst, 2004 ; Dragone and Raggi, 2021). An ideal analysis would also include cigarettes, a plainly addictive product, as in both Becker, Grossman, and Murphy (1994) and Dragone and Raggi (2021), unfortunately cigarette purchases are not available in the IRI data we have access to.

The Canonical Model , the AR(2) Model of Rational Addiction

Following Becker, Grossman, and Murphy (1994) we construct our AR(2) model of rational addiction. Consider a model with two goods and current utility in period t given by a concave utility function:

$$(1) \quad U(Y_t, C_t, C_{t-1}, e_t).$$

C_t is the quantity of the addictive good consumed in period t and C_{t-1} the quantity consumed in period $t-1$, Y_t is the consumption of a composite commodity in period t, and e_t reflects the impact of unmeasured life-cycle variables.

Individuals are assumed to be infinite-lived and to maximize total lifetime utility given discount rate r. Taking the composite commodity as numéraire, assuming the rate of interest is equal to the rate of time-preference (i.e $\beta = 1/(1+r)$), and denoting the price of the addictive commodity p_t the consumer's problem is:

$$(2) \quad max \sum_{t=1}^{\infty} \beta^{t-1} U(Y_t, C_t, C_{t-1}, e_t),$$

such that $C_0 = C^0$, the unobserved level of addictive consumption prior to the period under consideration and

$$\sum_{t=1}^{\infty} \beta^{t-1}(Y_t + P_t C_t) = A^0.$$

This specification ignores any effect of C on earnings and hence to the present value of wealth (A^0) as well as the effect of C on the length of life and all other types of uncertainty.

The associated first order conditions of this utility maximization problem are:

$$(3a) \quad U_y(C_t, C_{t-1}, Y_t, e_t) = \lambda,$$

$$(3b) \quad U_1(C_t, C_{t-1}, Y_t, e_t) + \delta U_2(C_t, C_{t-1}, Y_t, e_t) = \lambda P_t.$$

Equation (3a) is just the usual condition that the marginal utility of other consumption in each period is equal to the marginal utility of wealth, $U_y = \lambda$. Equation (3b) implies that the marginal utility of current consumption U_1, plus the discounted marginal effect on next periods utility U_{t+1}, equals the current price multiplied by the marginal utility of wealth. In the case of harmfully addictive products (e.g., cigarettes or liquor) U_2 is negative although this model assumes only that it is nonzero and the cases where $U_2 > 0$ indicate a beneficial addiction. When utility is not time-separable consumption depends not only on present prices but on prices in all periods through the effects of past and future prices on past and future consumption.

Consider a utility function which is quadratic in Y_t, C_t, and e_t. Solving the first order condition for Y_t and substituting into the first order condition for C_t the following model is obtained which forms the basis of our AR(2) analysis:

$$(4) \quad C_t = \theta C_{t-1} + \beta\theta C_{t+1} + \theta_1 P_t + \theta_2 e_t + \theta_3 e_{t+1},$$

where

$$\theta_1 = \frac{u_{yy}\lambda}{(u_{11}u_{yy} - u_{1y}^2) + \beta(u_{22}u_{yy} - u_{2y}^2)} < 0,$$

$$\theta_2 = \frac{-(u_{yy}u_{1e} - u_{1y}u_{ey})}{(u_{11}u_{yy} - u_{1y}^2) + \beta(u_{22}u_{yy} - u_{2y}^2)},$$

$$\theta_3 = \frac{-\beta(u_{yy}u_{2e} - u_{2y}u_{ey})}{(u_{11}u_{yy} - u_{1y}^2) + \beta(u_{22}u_{yy} - u_{2y}^2)},$$

where the lowercase letters denote the coefficients of the quadratic utility function, and the intercept is suppressed. θ_1 is negative by the concavity of U so equation (4) predicts a negative effect of price on current consumption, C_t where the marginal utilities of wealth, past consumption, and future consumption are fixed.

The effects of changes in past or future consumption on current consumption depend only on the term θ.

$$(5) \quad \theta = \frac{-(u_{12}u_{yy} - u_{12}u_{2y})}{(u_{11}u_{yy} - u_{1y}^2) + \delta(u_{22}u_{yy} - u_{2y}^2)} > 0,$$

When θ is positive past consumption reinforces current consumption, a necessary condition for addiction. A good is more addictive when the coefficient θ is larger (Becker, Grossman and Murphy, 1994).

Our estimated AR2 model follows from the specification in equation (4) following Becker, Grossman, and Murphy (1994) but accounting for the panel structure of our data. First using OLS, we estimate the following empirical specification of equation (4):

$$(6) \quad C_{it} = \alpha_0 + \alpha_1 C_{it-1} + \alpha_2 C_{it+1} + \alpha_3 P_{it} + \iota^{i} I_{it} + \tau_t + \mu_i + \epsilon_{it},$$

where $\alpha_1, \alpha_2, \alpha_3$ are equal to $\theta, \beta\theta, \theta_1$ respectively, notice that $\alpha_2/\alpha_1 = \beta$, the implied discount rate. Where consumption C is assumed to be equal to quarterly purchases of household i in quarter t. P_{it} is the county level price observed by household i in quarter t. The variable I_{it} is a categorical variable indicating the income group that household i belongs to in quarter t and thus each coefficient of ι^{i} indicates a control for a given income level. τ_t is a series of indicator variables for the year and quarter to control for seasonality and trend in consumption over time. μ_i is a household level fixed effect. In this model, a positive coefficient on α_1 is consistent with reinforcement and a positive coefficient on α_2 is consistent with forward looking behavior (Becker, Grossman, and Murphy, 1994; Chaloupka, 1991).

Notice that the specification of equation (4) presented in (6), the unobservables e_t and e_{t+1} are both represented in ϵ_{it}, the error term. These unobserved errors are likely to be serially correlated which would incorrectly imply that past and future consumption positively affect current consumption even when the true value of θ is zero (Becker, Grossman, and Murphy, 1994). This endogeneity problem is solved following from the suggestion in equation (4) that future and past prices affect current consumption only through future and past prices. Provided that the unobservables are uncorrelated with prices in these periods, past and future prices are logical instruments for C_{t-1} and C_{t+1}. Thus, we use 2SLS to estimate equation (6) using 3 lags and 3 leads of price as instruments for C_{t-1} and C_{t+1} following Becker, Grossman, and Murphy (1994). We implement this estimation with STATA's xtivreg command which implements G2SLS from Balestra and Varadharajan-Krishnakuma (1987). The results of these estimation procedures are presented in Tables 4 and 5 and discussed in the results section below.

A Recent Empirical Refinement, The AR(1) model of Rational Addiction

As has been previously noted, the AR(2) rational addiction process estimated in (6) has been shown to misidentify non-addictive products, like milk, as rationally addictive (Auld and Grootendorst, 2004). Thus, we also investigate an alternative specification of the rational addiction model proposed by Dragone and Raggi (2021) which they show to solve this problem using Auld and Grootendorst's (2004) data. We call AR(1) following their notation because this model only contains one auto-regressor for current consumption, past consumption. They observe that the Euler equation presented in equation (4) describes an infinity of candidate solutions. Among them only one is stationary, the saddle path. This saddle path solution to the optimal consumption path can be described by the following AR(1) equation:

$$(7) \quad C_t = \lambda C_{t-1} + \phi_1 P_{t-1} + \phi_2 P_t + \sum_{s=1}^{\infty} \phi_3(\mathrm{s}) P_{t+s} + \phi_0.$$

Based on the saddle path solution to (7) we modify the empirical specification of Dragone and Raggi (2021) to estimate a rational addiction model using household level panel data:

$$(8) \quad C_{it} = \alpha_0 + \alpha_1 C_{it-1} + \alpha_2 P_{it+1} + \alpha_3 P_{it} + \alpha_4 P_{it-1} + \iota^{i} I_{it} + \tau_t + \mu_i + \epsilon_{it}.$$

As in model (6) consumption C is assumed to be equal to quarterly purchases of household i in quarter t. P_{it} is the county level price observed by household i in quarter t. The variable I_{it} is a categorical variable indicating the income group that household i belongs to in quarter t and thus each coefficient of ι^{i} indicates a control for a given income level. τ_t is a series of indicator variables for the year and quarter to control for seasonality and trend in consumption over time. μ_i is a household level fixed effect. In this model the coefficient for α_1 must be positive to indicate reinforcement. However, the coefficient α_2 can be either positive, consistently with stockpiling behavior, or negative, indicating that consumption today decreases in expectation of a future price or tax increase (Dragone and Raggi, 2021).

In equation (8) lagged consumption may also be endogenously related to the error term. Therefore, Dragone and Raggi (2021) recommend instrumenting for C_{it-1} using P_{t-2} and P_{t+2} to deal with this concern. While Dragone and Raggi (2021) find that endogeneity concerns due to the presence of lagged consumption do not pose a relevant threat for their empirical estimation we use a 2SLS estimation procedure like we used in the AR(2) specification to investigate this potential endogeneity bias.

Using the models described by equations (6) and (8) we empirically investigate IRI consumer panel and retail scanner data. Additionally, to account for bias in standard deviation due to individual households making multiple decisions we estimate all of these models with cluster standard errors. The results of these estimation procedures are presented in Tables 7 and 8 and discussed in the results section below.

A Novel Strategy, Random Coefficients Regression

We present a novel (to the best of our knowledge) empirical strategy to modelling rational addiction which takes advantage of the micro scale of our data to assess the heterogeneity of addictive behavior within a population. Here we estimate a random coefficients model which is a natural extension of both the canonical AR(2) model and the novel AR(1) model discussed above. It can be helpful to think of this class of models as empirical refinements on the ad-hoc strategy of estimating individual regression results for each of the households and summarizing the means and standard deviations of their parameter estimates. This methodology simply proposes an asymptotically efficient estimator for the mean coefficients vector and an unbiased estimator for the variance-covariance matrix (Swamy, 1971).

In equations (9) and (10) we present derivations of these models for the AR(2) model, as described in equation (6), and the AR(1) model, as described in equation (8), respectively following from Green (2012) chapter 11:

$$(9) \quad C_{it} = \alpha_{i0} + \alpha_{i1} C_{it-1} + \alpha_{i2} C_{it+1} + \alpha_{i3} P_{it} + \epsilon_i,$$

$$(10) \quad C_{it} = \alpha_{i0} + \alpha_{i1} C_{it-1} + \alpha_{i2} P_{it+1} + \alpha_{i3} P_{it} + \alpha_{i4} P_{it-1} + \epsilon_i,$$

$$E[\epsilon_i | X_i] = 0,$$

$$E[\epsilon_i \epsilon_i' | X_i] = \sigma^2,$$

where

$$\alpha_{ij} = \alpha_j + u_{ij}, \qquad j \in 1,2,3,4,$$

and

$$E[u_i | X_i] = 0,$$

$$E[u_i u_i' | X_i] = \Gamma.$$

This approach assumes no autocorrelation or cross-correlation in ϵ_i, this assumption is clearly problematic, as shown above autocorrelation is likely in both models of rational addiction and will be addressed in later iterations of this work. It is also necessary that for each household the number of observations exceeds the number of parameters to be estimated so that is possible to estimate a linear regression of C_{ti} on the model specific covariates X_i. In practice, we find that limiting the sample to only households which consume the product in question more at least 4 times a year is sufficient to meet this condition.

The α_{ij} which applies to a particular household is the outcome of a random process with mean vector α_j and covariance matrix Γ. The generalized regression model for each block of household level observations is thus:

$$y_i = X_i \alpha_j + (\epsilon_i + X_i u_i),$$

so

$$\Omega_{ij} = E\left[\left(C_{ti} - X_i \alpha_j \right) \left(C_{ti} - X_i \alpha_j \right)' \middle| X_i \right] = \sigma_\epsilon^2 I_T + X_i \Gamma X_i'.$$

For the system as a whole, the disturbance matrix is block diagonal, with $T\ x\ T$ diagonal block Ω_{ij}. We can write the GLS estimator as a matrix weighted averaged of the household specific OLS estimators.

$$\hat{\alpha_J} = (X'\Omega X)^{-1} X' \Omega^{-1} C_t = \sum_{i=1}^{n} W_i a_{ij},$$

where

$$W_i = [\sum_{i=1}^{n} (\Gamma + \sigma_\epsilon^2 (X_i' X_i)^{-1})^{-1}]^{-1} (\Gamma + \sigma_\epsilon^2 (X_i' X_i)^{-1})^{-1}.$$

Empirical estimation of this model requires an estimator of Γ. We utilize the approach of Swamy (1970) which uses the empirical variance of the set of n household-level least squares estimates, a_i minus the average values of $s_i^2 (X_i' X_i)^{-1}$:

$$G = \left[\frac{1}{n-1} \right] [\Sigma_i a_i a_i' - n \bar{a} \bar{a}'] - \left(\frac{1}{N} \right) \Sigma_i V_i,$$

where

$$\bar{a} = \left(\frac{1}{n}\right) \Sigma_i a_i$$

and

$$V_i = s_i^2 (X_i' X_i)^{-1}$$

We estimate these models using the xtrc function in STATA 17 (2021). While the average parameter estimates are interesting and are discussed in tables 6 and 7 we also make use of the individual household level regressions. Specifically, we identify if individual households exhibit rational addiction to either regular soda or milk defining rational addiction as a negative price coefficient, significant (at the 5% level) positive effect of past consumption on current consumption, and a significant (at the 5% level) effect of either future consumption or future prices on current consumption in the case of the AR(2) and AR(1) models respectively. The results of this estimation are displayed in tables 6 and 9 and discussed in the results section below.

Dynamic Structural Model of Addiction and Stockpiling

A final approach which we intend to implement follows from Gordon and Sun's (2015) dynamic structural model of addiction and stockpiling. This analysis addresses a crucial assumption of our previous analysis, assuming that purchases are reflective of consumption. However, consumers can stockpile past purchases, in the case of soda for a large span of time. Addiction and stockpiling mechanisms suggest opposite effects of past purchases on current purchases. Where addiction implies that past consumption increases present consumption, stockpiling implies that past purchases decrease present purchases. Therefore, not accounting for stockpiling could cause us to underestimate addictive behavior in soda relative to a less storable product, such as milk. This modelling strategy is complex and computationally intensive, so we have not yet implemented it. However, much of the complex data cleaning to implement this estimation strategy is now complete and we intend to utilize this strategy in the near future.

Gordon and Sun (2015) present a model of addiction which combines dynamic model endogenous consumption, as developed in Hendel and Nevo (2006a), with standard models of rational addiction. Addiction can be thought of as a stock a_{it} similar to inventory $I_{i,t}$, which when combined with prices P_t comprise the state space $s_{i,t}$ of individual i in time t. Addiction follows a law of motion where it increases with consumption $c_{i,t}$, but the stock of addiction depreciates at rate δ_i:

$$(9) \quad a_{i,t+1} = (1 - \delta_i) a_{i.t} + c_{i,t}$$

Inventory follows a similar law of motion decreasing in consumption but increasing in purchases with d_{itjq} being a dummy variable describing the decision to purchase quantity q and product-tier combination j:

$$(10) \quad I_{i,t+1} = I_{i,t} + \sum_{j,q} d_{itjq}\, q_{itjq} - c_{it}$$

With these two stocks described the consumer can be described as solving an infinite time horizon dynamic programming problem choosing consumption and making purchase decisions. The per period utility of individual i in period t can thus be described as:

$$(11) \quad U\left(c_{it}, d_{it}, s_{i,t}; \theta_i\right) = u_c\left(c_{i,t}, a_{i,t}; \alpha_i\right) + u_p\left(d_{i,t}, P_t; \beta_i, \xi_i\right) - C\left(I_{i,t}; h\right)$$

Where u_c is the utility of consumption, a quadratic function of addiction and consumption. u_p is purchase utility basically measuring disutility of spending money as a function of price sensitivity and spending controlling for consumer fixed effects. $C\left(I_{i,t}; h\right)$ is just the cost of holding inventory, which increases linearly as the inventory increases. They also must calculate the expectation of future prices as well as the probability of store visits given that a store visit is observed in the previous period $\rho_{1,t}$ or not ρ_{i0}.

Given consumers' current state period utility as well as their expectations of future store visits consumers simultaneously make consumption decisions c_{it}^* and purchase decisions to solve an infinite time horizon dynamic programming problem. This can be represented using Bellman equations assuming a known and fixed discount factor $\beta = 0.995$.

For the period in which a purchase is observed:

$$(12) \quad V(s_{it}) = max_{c_{it}d_{it}}\{U(c_{it}, d_{it}, s_{it}; \theta) + \beta\mathbb{E}[\rho_{i1}\, V(s_{it+1} \;+ (1 - \rho_{i1}W(s_{it+1})|s_{it}]\}$$

$$s.t \quad 0 \leq c_{it} \leq I_{it} + \sum d_{itjq}q_{itjq}$$

For periods when no purchase is observed:

$$(13) \quad W(s_{it}) = \max_{c_{it}d_{it}}\{u_c(c_{i,t}, a_{i,t}; \alpha_i) - C(I_{it}\,; h_t) + \beta\mathbb{E}[(1 - \rho_{i0})\, V(s_{it+1}) + \rho_{i0}W(s_{it+1})|s_{it}]\}^2$$

$$s.t\; 0 \leq c_{it} \leq I_{it}$$

Purchases d_{it} are observed in the data so the only unknown that is solved for is consumption by maximizing $V(s_{it})|d_{it}$. The probability that the individual makes the observed purchase decision $D_{i,t}$ is expressed as a logit choice probability assuming the error term is distributed i.i.d. extreme value:

$$(14) \quad Pr\left(D_{it} = d_{itjq}\middle|a_{it}, I_{it}; \theta_m\right)\frac{exp(V_{itjq}^m(s_{it}; \theta_m))}{\Sigma_{j'q'}\, expV_{itjq}^m(s_{it}; \theta_m)}$$

This probability can be aggregated up through consumers and consumer types (m describing light and heavy use consumers) to describe a log likelihood function over all households. This function can be maximized using a nested fixed-point method like that discussed in Hendel and Nevo (2006a) to estimate the most likely state of all consumers in each time period including their addiction parameters and inventories.

We intend to implement this model using our data, adding to the food addiction literature as well as the literature concerning health impacts of addiction. First, this modelling strategy allows us to de-couple the effects of stockpiling, which is suggested to be a present dynamic in table 3, from the effects of addiction. Addiction stockpiling are two dynamic forces which work in opposite directions, addiction suggests that if consumers have bought something in the past they

---

[2] There is a typo in Gordon and Sun's methodology relating to the conditional probability, I have changed this equation to match what they intend to describe, and I assume they do describe.

are likely to buy more of it in the future, stockpiling on the other hand suggests that past purchases limit current purchases. In our reduced-form models we assume that monthly purchases are equal to monthly consumption, in the case of soda especially which can be easily stored for long periods of time disaggregating the dynamics of addiction and stockpiling is necessary. Furthermore, it has been suggested that addiction is related to other health conditions, as we have access to med-profiler data we can integrate a consumer's health state into their state space. Some interesting measures of physical and mental health we can use for this purpose are BMI/obesity, depression, anxiety, and whether the individual is currently quitting smoking. Accounting for these and potentially other health factors from an economics point of view can be a valuable contribution to the addiction literature.

V. Results

Descriptive Analysis

First we replicate Gordon and Sun's (2015) descriptive evidence of addiction in table 2. Here we see the proportion of weekly purchases where the household's purchase quantity increased, decreased, or stayed the same. If a consumer was addicted to a product they would be more likely to increase their purchases week to week. Households are slightly more likely to increase their purchase quantity of regular soda than decrease it whereas, in the case of milk, consumers are about equally likely to increase or decrease their purchase quantities week to week. This is rather weak evidence that regular soda is addictive and low-fat/skim milk is not.

We then replicate descriptive analysis of stockpiling presented in Hendel and Nevo (2006b). Following Hendel and Nevo (2006b), we define a sale as a week where the customers modal brand and size at the stores they visit sells for 5% lower than its modal price at that store. We show that for both products, when there is a sale, consumers buy a larger amount on sale for a lower per volume price and that there is a lower number of weeks since the last purchase than weeks until the next purchase, i.e., consumers buy earlier and buy more so their stockpile lasts longer. However, milk is bought much more frequently than soda, likely in part due to soda's increased storability. This result suggests that stockpiling is a present dynamic in the demand for both products.

AR(2) Rational Addiction Models

Table 4 presents the results of estimating equation (6), the AR(2) model of Becker, Grossman, and Murphy (1994). In this class of models, rational addiction is defined as a significant positive effect of past consumption on current consumption, as well as a significant effect of future consumption on current consumption. We first perform these estimates using OLS and we see evidence of rational addiction to soda. However, as in Auld and Grootendorst (2004), we see that this model also predicts rational addiction for low-fat/skim milk. Turning now to the 2SLS estimates, implementing the instrumental variable strategy of using three lags and three leads of price to instrument for endogenous lead and lag consumption. This method still shows significant evidence that low-fat/skim milk is rationally addictive. Curiously though, the coefficient for the effect of lag consumption on present consumption for regular soda has turned from positive to negative, this might imply that, after controlling for the endogenous relationship between past and present consumption due to serial autocorrelation, that stockpiling dynamics override addiction dynamics. As mentioned above, stockpiling presents as the opposite dynamic to addiction, where addiction implies that past consumption should increase present consumption, stockpiling implies that past purchases should decrease present purchases. It is not unreasonable to think that stockpiling dynamics are stronger for soda than milk, after all soda can be stored effectively indefinitely without the need for refrigeration.

Table 5 shows results for the AR(2) estimation with the sample split by BMI group. These results confirm again that the milk addiction paradox is a present problem in this modelling strategy. It is interesting to note however that the OLS estimator predicts rational addiction to soda to be significantly stronger in the obese group than the other groups, in terms of the magnitude of the coefficients of lag and lead consumption. In the 2SLS model while it is likely that the stockpiling dynamic is bringing down the coefficient of lag consumption there is still the most evidence for rational addiction to soda in the obese group. There is not much pattern in the strength of rational addiction across BMI groups for low-fat/skim milk.

Finally, in table 6, we show the results of our novel random coefficients model of rational addiction, which is discussed in detail in the methodology section of this paper. Overall, this modelling strategy confirms many of the results of the previous modelling strategy; the AR(2) model consistently shows that an unreasonably large portion of households, 30.2% are rationally addicted to low-fat/skim milk. While an equally unbelievably large but surprisingly smaller proportion of households, 21.4%, show evidence of rational addiction to regular soda. Once again however, we see an interesting pattern with respect to average household obesity. Obese households are much more likely to be classified as rationally addicted to soda than non-obese households. While this pattern is true for milk as well the magnitude of the difference is smaller. Our novel random coefficients modelling strategy should better capture heterogeneity in addictive behavior but does not solve the primary problem of AR(2) models of rational addiction, that they produce unrealistic estimates of plainly non-addictive products, like milk.

We confirm the result of Auld and Grootendorst (2004), that AR(2) models of rational addiction are not particularly useful because they show addiction to non-addictive products. However, we also show evidence that stockpiling dynamics are present in the demand for regular soda, so Gordon and Sun's (2017) model of addiction and stockpiling may be effective in disaggregating stockpiling and addiction dynamics. Additionally, we see evidence that rational addiction to soda is stronger for obese consumers than non-obese consumers.

AR(1) Rational Addiction Models

Now we turn to the recent empirical refinement to rational addiction models, the AR(1) model of Dragone and Raggi described in detail in the methodology section above. Dragone and Raggi purport that this class of models should solve the milk addiction paradox, i.e., that they should not show rational addiction to non-addictive products. Identifying rational addiction in this model relies on observing a reinforcement effect, a significant positive effect of past consumption on current consumption, and forward-looking behavior, in this class of models that takes the form of a significant effect of future prices on current consumption. While this class of models seems to have much more realistic predictions regarding the addictiveness of milk and soda, there are still some empirical irregularities that are worth highlighting.

First, we show results of estimating equation (8) using the full sample. These results are significantly more believable than the AR(2) results. Here we see that under no model is either regular soda or low-fat/skim milk classified as rationally addictive. The 2SLS results, instrumenting for past consumption using the second lag and second lead of price, show some interesting differences, even though in Dragone and Raggi's (2021) estimation this instrumentation strategy is not hugely influential on the results. First, we see that for regular soda the 2SLS estimates imply a negative relationship between past consumption and current consumption, this could be indicative of stockpiling dynamics as noted with the AR(2) models as well. There is also a rather strange result in the 2SLS estimation of this model for low-fat/skim milk, in this estimation the coefficient estimate for past consumption nearly doubles in magnitude. Rational addiction models are usually primarily concerned with significance and indeed as the coefficient on future price is not significant this product is “not rationally addictive”. Nevertheless, assuming that this instrumentation strategy is effective, this result implies that milk consumption has a strong reinforcement effect. This is not the only strange result we observe for low-fat/skim milk when estimating this model using the AR(1) methodology.

In table 8 we run our AR(1) analysis by sub-samples of household average obesity class. For regular soda, these results generally confirm some of our previous assessments. For a start, we do not show evidence of rational addiction in any BMI group. However, there is arguably stronger evidence of rational addiction in the obese group with larger positive estimates of the effect of past consumption on present consumption in both the OLS and 2SLS estimates. For low-fat/skim milk we do see a surprising result, however. The OLS estimator predicts rational addiction to low-fat/skim milk in both the normal and overweight BMI groups, there is a significant positive coefficient of past consumption and a significant negative coefficient for future price implying both reinforcement and forward-looking behavior. Using the 2SLS estimator we no longer find evidence of rational addiction; however, we again see very high estimates of the effect of past consumption on present consumption. This strange result merits further exploration and a thorough re-check of our code to ensure that this is not a mistake. However, assuming that it is correctly estimated, this result implies that Dragone and Raggi’s (2021) empirical refinement on the rational addiction model may not wholly solve the milk addiction paradox.

We come now to our preferred specification, this model combines the AR(1) model, which has shown to provide much more realistic predictions than the AR(2) model with our novel random coefficients estimation strategy, which more accurately models addiction as a heterogeneously distributed phenomenon within a population. In table 9 we show average coefficient estimates for our random coefficients estimation strategy which, once again, is discussed in detail in the methodology section of this paper. This estimation method produces the most realistic predictions of rational addiction to both regular soda and low-fat/skim milk. First, regarding regular soda we do not see strong evidence of rational addiction in the aggregate results. Looking at the household level regressions we see that only about 1% of households are rationally addicted to regular soda, based on the definition of a significant (at the 5% level) positive effect of past consumption and a significant (again at the 5% level) effect of future price on current consumption. This result is especially weak when we consider that if we assumed that the coefficients of past consumption and future price had zero mean and were completely independent we would still predict that 0.25% of consumers are rationally addicted to any product because of the nature of our 5% significance cutoffs. However, it is interesting to note that the proportion of households that show rational addiction to soda is higher among obese households (1.21%) than households with an average BMI in the normal or overweight BMI groups (0.97% and 0.85% respectively). While the aggregate results for low-fat/skim milk suggest that rational addiction is a significant dynamic in the consumption of low-fat/skim milk the individual level results tell another story. Only 2.34% of households are rationally addicted to low-fat/skim milk, while this is a higher portion of households than for regular soda it is hardly substantial. Although, it likely explains the more significant result for the future consumption coefficient on average. Overall, this model predicts very small portions of the population are addicted to either regular soda or low-fat/skim milk. This is also without a strategy to account for serial correlation in past consumption as with the instrumental variable strategy of previous estimation methods, which we intend to explore in a future iteration of this paper.

AR(1) models generally show very little evidence of rational addiction to both low-fat/skim milk and regular soda. This makes sense as both of these products are common grocery items and not, say, heroin. However, it is likely that our modelling strategy could use some improvement. For one, there are a couple of strange results when estimating AR(1) rational addiction models for low-fat/skim milk that may suggest that Dragone and Raggi's (2021) empirical refinement is not completely effective in resolving the milk addiction paradox. Furthermore, stockpiling dynamics, having the opposite implications for the effect of past purchases on current purchases as addiction, could be obscuring addiction dynamics, especially in the case of regular soda which is very storable. Future research on this topic will implement Gordon and Sun's (2016) dynamic model of addiction and stockpiling to address this concern.

VI. Discussion

The results of this paper suggest that rational addiction is not an effective framework for assessing the demand for regular soda. While the canonical model of rational addiction of Becker, Grossman, and Murphy (1994), which is commonly employed in the literature, shows strong evidence of rational addiction to regular soda it also shows strong evidence of rational addiction to low-fat/skim milk, this problem is well known in the literature (Auld and Grootendorst, 2004). However, using an updated model of Dragone and Raggi (2021) which we show is mostly effective in dealing with this "milk addiction paradox" we find little evidence of rational addiction to regular soda. Our preferred modelling strategy, a novel random coefficients treatment of rational addiction applied to this AR(1) model, shows that only 1% of households show evidence of rational addiction to regular soda.

Nevertheless, this paper does show some interesting patterns with respect to addiction that merit further exploration. First, we consistently show stronger evidence of rational addiction to regular soda in households whose BMI classifies them as obese. Furthermore, we show evidence that stockpiling dynamics are likely present, especially in the demand for regular soda. Stockpiling likely obscures evidence of rational addiction as the two processes suggest opposite dynamics with respect to the effect of past purchases on current purchases. For this reason, we intend to explore Gordon and Sun's (2015) dynamic model of addiction and stockpiling in later iterations of this work, indeed much of the groundwork for this analysis has been established in the completion of this paper.

Overall, this paper adds to the robust modern literature on rational addiction (e.g., Dragone and Raggi, 2021;Picoli and Tiezzi, 2021; Laporte, Dass, and Ferguson, 2017; Gordon and Sun, 2016) in a couple of important ways. First, we present some of the most empirically robust analysis of food addiction using modern scanner data to date, Richards, Patterson, and Tegene (2007) had access to data for only 30 households where we present analysis of IRI consumer panel data for thousands of households. Second, we explore health implications with respect to BMI of rational addiction results taking advantage IRI's underutilized Medprofiler survey. Finally, we take advantage of our large micro-scale dataset to present a novel random coefficients strategy for estimation of rational addiction models. This strategy more accurately represents addiction as a phenomenon which is heterogeneously distributed within households.

Works Cited

Auld, M. Christopher, and Paul Grootendorst. 2004. "An Empirical Analysis of Milk Addiction." *Journal of Health Economics* 23 (6): 1117–33. https://doi.org/10.1016/j.jhealeco.2004.02.003.

Baltagi, Badi H., and James M. Griffin. 2002. "Rational Addiction to Alcohol: Panel Data Analysis of Liquor Consumption." *Health Economics* 11 (6): 485–91. https://doi.org/10.1002/hec.748.

Balestra, Pietro, and Jayalakshmi Varadharajan-Krishnakumar. 1987. "Full Information Estimations of a System of Simultaneous Equations with Error Component Structure." *Econometric Theory* 3 (2): 223–46. https://doi.org/10.1017/S0266466600010318.

Becker, Gary S, Michael Grossman, and Kevin M Murphy. 1994. "An Empirical Analysis of Cigarette Addiction." *American Economic Association* 84 (3): 396–418.

Becker, Gary S, and Kevin M Murphy. 1988. "A Theory of Rational Addiction Author (s): Gary S . Becker and Kevin M . Murphy Published by : The University of Chicago Press Stable URL : Http://Www.Jstor.Org/Stable/1830469 REFERENCES Linked References Are Available on JSTOR for This Article : You Ma." *Journal of Political Economy* 96 (4): 675–700. https://www.jstor.org/stable/pdf/1830469.pdf?refreqid=excelsior%3A9a5a68a3dc5a8438d77c84cf9e4b860c.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–90.

Cameron, Samuel. 1999. "Rational Addiction and the Demand for Cinema." *Applied Economics Letters* 6 (9): 617–20. https://doi.org/10.1080/135048599352736.

Castiglione, Concetta, and Davide Infante. 2016. "Rational Addiction and Cultural Goods: The Case of the Italian Theatregoer." *Journal of Cultural Economics* 40 (2): 163–90. https://doi.org/10.1007/s10824-015-9247-4.

Cawley, John. 2004. "An Economic Framework for Understanding Physical Activity and Eating Behaviors." *American Journal of Preventive Medicine* 27 (3 SUPPL.): 117–25. https://doi.org/10.1016/j.amepre.2004.06.012.

Chaloupka, Frank. 1991. "Rational Addictive Behavior and Cigarette Smoking." *The University of Chicago Press Stable* 99 (4): 722–42. http://www.jstor.com/stable/2937778

Chen, Xuan, Yizao Liu, Edward C. Jaenicke, and Adam N. Rabinowitz. 2019. "New Concerns on Caffeine Consumption and the Impact of Potential Regulations: The Case of Energy Drinks." *Food Policy* 87 (August 2018): 101746. https://doi.org/10.1016/j.foodpol.2019.101746.

Dockner, Engelbert, and Gustav Feichtinger. 1993. "Cyclical Consumption Patterns and Rational Addiction Author (s): Engelbert J . Dockner and Gustav Feichtinger Source : The American Economic Review , Vol . 83 , No . 1 (Mar ., 1993), Pp . 256-263 Published by : American E." *The American Economic Review* 83 (1): 256–63.

Dragone, Davide, and Davide Raggi. 2021. "Resolving the Milk Addiction Paradox." *Journal of Health Economics* 77: 102452. https://doi.org/10.1016/j.jhealeco.2021.102452.

Dragone, Davide, and Davide Raggi. 2018. "Testing Rational Addiction: When Lifetime Is Uncertain, One Lag Is Enough." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3155099.

Greene, W. H. 2012. Econometric Analysis. 7th ed. Upper Saddle River, NJ: Prentice Hall.

Gordon, Brett R., and Baohong Sun. 2015. "A Dynamic Model of Rational Addiction: Evaluating Cigarette Taxes." *Marketing Science* 34 (3): 452–70. https://doi.org/10.1287/mksc.2014.0885.

Grossman, Michael, and Frank J. Chaloupka. 1998. "The Demand for Cocaine by Young Adults: A Rational Addiction Approach." *Journal of Health Economics* 17 (4): 427–74. https://doi.org/10.1016/S0167-6296(97)00046-5.

Hendel, Igal, and Aviv Nevo. 2006. "Sales and Consumer Inventory." *RAND Journal of Economics* 37 (3): 543–61. https://doi.org/10.1111/j.1756-2171.2006.tb00030.x.

Hendel, Igal, and Aviv Nevo. 2006. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica* 74 (6): 1637–73. www.jstor.org/stable/4123086.

Laporte, Audrey, Adrian Rohit Dass, and Brian S. Ferguson. 2017. "Is the Rational Addiction Model Inherently Impossible to Estimate?" *Journal of Health Economics* 54: 161–75. https://doi.org/10.1016/j.jhealeco.2016.12.005.

Liu, Jin Long, Jin Tan Liu, James K. Hammitt, and Shin Yi Chou. 1999. "The Price Elasticity of Opium in Taiwan, 1914-1942." *Journal of Health Economics* 18 (6): 795–810. https://doi.org/10.1016/S0167-6296(99)00023-5.

Liu, Xiaoou, and Rigoberto Lopez. 2012. "Evidence of Rational Addiction to Carbonated Soft Drinks?" *China Agricultural Economic Review* 4 (3): 300–317. https://doi.org/10.1108/17561371211263347.

Lenoir, M., Serre, F., Cantin, L., & Ahmed, S. H. (2007). Intense sweetness surpasses cocaine reward. PloS one, 2(8), e698. https://doi.org/10.1371/journal.pone.0000698

Olekans, Nils and Peter Bardsley. 1996. "Rational Addiction to Caffeine : An Analysis of Coffee Consumption Reproduced with Permission of the Copyright Owner . Further Reproduction Prohibited without Permission ." *Journal of Political Economy* 104 (51): 1100–1104.

Piccoli, Luca, and Silvia Tiezzi. 2020. “Rational Addiction and Time Consistency: An Empirical Test.” *Journal of Health Economics* 80. https://doi.org/10.1016/j.jhealeco.2021.102546.

Richards, Timothy J., Paul M. Patterson, and Abebayehu Tegene. 2007. “Obesity and Nutrient Consumption: A Rational Addiction?” *Contemporary Economic Policy* 25 (3): 309–24. https://doi.org/10.1111/j.1465-7287.2007.00047.x.

Richards, Timothy J., and Stephen F. Hamilton. “Obesity and hyperbolic discounting: an experimental analysis.” *Journal of Agricultural and Resource Economics* (2012): 181-198.

Rogeberg, Ole. 2004. “Taking Absurd Theories Seriously: Economics and the Case of Rational Addiction Theories*.” *Philosophy of Science* 71 (3): 263–285. https://www.jstor.org/stable/10.1086/421535

Roodman, David. 2009. “How to Do Xtabond2: An Introduction to Difference and System GMM in Stata.” *Stata Journal* 9 (1): 86–136. https://doi.org/10.1177/1536867x0900900106.

Swamy, P.A.V.B. 1970. “Efficient Inference in a Random Coefficient Regression Model”. *Econometrica* 38 (2) http://links.jstor.org/sici?sici=0012-9682%28197003%2938%3A2%3C311%3AEIIARC%3E2.0.CO%3B2-A

World Bank. 2005. “Introduction to Poverty Analyses.” *Poverty Manual*, no. August: 1–218. http://siteresources.worldbank.org/PGLP/Resources/PovertyManual.pdf.

Zhen, C., Finkelstein, E. A., Nonnemaker, J. M., Karns, S. A., & Todd, J. E. “Predicting the effects of sugar-sweetened beverage taxes on food and beverage demand in a large demand system.” *American Journal of Agricultural Economics* 96.1(2013): 1-25.

Additional Sources

Alston, Julian M., and Abigail M. Okrent. "US Farm Subsidies and Obesity the Effects of Farm and Food Policy on Obesity in the United States”. *Palgrave Macmillan, New York*, 2017.Economic Research Service (ERS). “Using Scanner Data”. *USDA*. (January 25, 2021) https://www.ers.usda.gov/topics/food-markets-prices/food-prices-expenditures-costs/using-scanner-data/

Murray, Emily. “Alcohol Addiction and Abuse.” *Addiction Center,* Addiction Center, 8 Dec.

2021, https://www.addictioncenter.com/alcohol/.

Ogden C.L., Carroll M.D., Kit B.K., and Flegal K.M. “Prevalence of Childhood and Adult Obesity in the United States, 2011-2012.” *Journal of the American Medical Association* 311(2015): 806-814.

StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC.

United States Center for Disease Control and Prevention (CDC). "Assessing Your Weight".

(September 17, 2020). United States Center for Disease Control and Prevention (CDC). "Obesity is a common, serious, and costly disease". (June 7, 2021).
https://www.cdc.gov/obesity/data/adult.html

Appendix

Table 1. Summary Statistics by Product

	Months in Sample	Average Price per fl.oz	Quantity Purchased (fl.oz)	Average BMI	Households Size	Households with Children	Number of Households
Regular Soda	65.21 (8.50)	¢3.06 (¢2.05)	215.77 (352.66)	29.40 (6.04)	2.29 (1.12)	26.90%	5,437
Low-fat /Skim Milk	65.73 (8.23)	¢2.97 (¢1.64)	157.85 (167.57)	28.93 (5.69)	2.13 (1.00)	19.92%	9,315

Note: Mean quantities for each product are reported with standard deviations in parentheses.

Table 2. Descriptive Analysis of Addiction

	Regular Soda	Low Fat/Skim Milk
Weeks with same Purchase Quantity as Last Week	5.5%	17.8%
Weeks with Purchase Quantity Increase Since Last Week	47.5%	41.1%
Weeks with Purchase Quantity Decrease Since Last Week	47.0%	41.1%
Difference	0.5%	0.04%
Z-Stat	8.31	0.01

Notes: The quantities in the first three rows correspond to the probability that a consumer purchases the same, bigger, or smaller quantities on the current purchase occasion compared to the previous purchase occasion. The row "z-stat" reports the test statistic under the null hypothesis that "increasing" equals "decreasing" and the alternative that "increasing" > "decreasing."

Table 3. Descriptive Analysis of Stockpiling						
	Regular Soda			**Low-Fat/Skim Milk**		
	Off Sale	On Sale	T-Stat	Off Sale	On Sale	T-Stat
Weekly Quantity Units	3.51	3.99	29.88	2.60	2.71	19.60
	(4.38)	(4.82)		(2.37)	(2.29)	
Price per Unit	$2.58	$2.57	-1.85	$2.75	$2.51	-114.76
	($2.03)	($1.54)		($0.99)	($0.83)	
Weekly Fluid Ounces	640.35	809.25	54.16	455.31	499.54	44.42
	(857.05)	(936.02)		(417.10)	(438.99)	
Price per Fluid Ounce	$0.02	$0.02	-74.42	$0.02	$0.02	-118.60
	($0.02)	($0.02)		($0.02)	($0.01)	
Weeks Since Last Purchase	10.69	10.93	2.45	5.79	4.74	-26.16
	(28.20)	(28.59)		(18.36)	(15.54)	
Weeks Until Next Purchase	10.52	11.22	6.99	5.46	5.04	-10.74
	(28.10)	(28.76)		(17.64)	(16.29)	
Difference in Difference Between Weeks Since Last Purchase and Weeks Until Next Purchase						
Difference	0.17	-0.29		0.32	-0.29	
T-Stat	1.95	-2.52		7.44	-8.03	

Notes: The “Off-Sale” column reports the mean quantities for each row restricted to purchase observation that occurred when the chosen item was not on sale. The "On-Sale" reports the reports the mean quantities for each row restricted to purchase observation that occurred when the chosen item was on sale. Standard errors are in parentheses. T-statistics are reported for difference in means tests between off-sale and on-sale values and for the difference between the weeks since last purchase and weeks until next purchase.

Table 4. Estimates of AR(2) Model of Rational Addiction

	Regular Soda		Low-fat Milk	
	OLS	2SLS	OLS	2SLS
C_{t-1}	0.142*** (0.0180)	-0.298* (0.143)	0.267*** (0.00526)	0.166** (0.0520)
C_{t+1}	0.139*** (0.0160)	0.171* (0.0829)	0.268*** (0.00522)	0.258*** (0.0736)
P_t	-515.8*** (47.47)	-730.8*** (133.4)	-141.8 (101.7)	-152.5 (101.8)
Rational Addiction	Yes	No	Yes	Yes

Notes: The dependent variable in these models is present consumption. All models include individual level fixed effects, year and month factor variables, and controls for household income. Standard errors are clustered at the household level. Instruments for the 2SLS estimator include 3 lag and 3 lead prices, following Becker and Murphy (1994). In the AR2 model rational addiction is shown by a significant positive coefficient on C_{t-1} and a significant coefficient on C_{t+1}.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 5. Estimates of AR(2) Models of Rational Addiction by BMI Group

	OLS			2SLS		
Regular Soda	Normal	Overweight	Obese	Normal	Overweight	Obese
C_{t-1}	0.0940*** (0.0238)	0.0719*** (0.00937)	0.174*** (0.0326)	-0.445*** (0.121)	-0.0367 (0.283)	-0.0964 (0.110)
C_{t+1}	0.0899*** (0.0227)	0.0732*** (0.00920)	0.167*** (0.0286)	0.0106 (0.173)	0.0651 (0.149)	0.134 (0.100)
P_t	-509.5*** (122.4)	-434.8*** (77.51)	-663.0*** (80.21)	-636.9** (232.6)	-508.8** (154.8)	-877.5*** (103.3)
Rational Addiction	Yes	Yes	Yes	No	No	No

	OLS			2SLS		
Low-Fat/Skim Milk	Normal	Overweight	Obese	Normal	Overweight	Obese
C_{t-1}	0.247*** (0.0123)	0.224*** (0.00663)	0.262*** (0.0102)	0.206 (0.107)	0.288*** (0.0710)	0.193 (0.107)
C_{t+1}	0.250*** (0.0125)	0.227*** (0.00688)	0.262*** (0.0101)	0.453*** (0.114)	0.355*** (0.0666)	0.264* (0.108)
P_t	-509.5*** (122.4)	-434.8*** (77.51)	-663.0*** (80.21)	-444.8*** (87.76)	-472.9*** (59.70)	-67.74 (55.11)
Rational Addiction	Yes	Yes	Yes	Yes	Yes	No

Notes: The dependent variable in these models is present consumption. All models include individual level fixed effects, year and month factor variables, controls for household income, and a variable for lagged price. Standard errors are clustered at the individual level. Instruments for the 2SLS estimator include P_{t-2} and P_{t+2}, following Dragone and Raggi(2021).In the AR1 model rational addiction is shown by a significant positive coefficient on C_{t-1} and a significant coefficient on P_{t+1}.

* p<0.05, ** p<0.01, *** p<0.001

Table 6. Average Results for Random Coefficients Estimation of AR(2) models of Rational Addiction

	Regular Soda	Low-fat/Skim Milk
C_{t-1}	0.072*** (0.003)	0.146*** (0.002)
C_{t+1}	0.072*** (0.003)	0.148*** (0.002)
P_t	-2046.38*** (245.66)	-273.66*** (82.51)
Proportion of Households with Rational Addiction	21.4%	30.2%
Proportion of Normal BMI Households with Rational Addiction	19.9%	29.3%
Proportion of Overweight BMI Households with Rational Addiction	20.0%	29.2%
Proportion of Obese BMI Households with Rational Addiction	23.7%	31.2%

Notes: The dependent variable in these models is present consumption. Estimation procedures for parameters in this model are described in the methodology section. In the AR2 model rational addiction is shown by a significant positive coefficient on C_{t-1} and a significant coefficient on C_{t+1}.

* p<0.05, ** p<0.01, *** p<0.001

Table 7. Estimates of AR(1) Model of Rational Addiction

	Regular Soda		Low-fat Milk	
	OLS	2SLS	OLS	2SLS
C_{t-1}	0.170*** (0.0231)	-0.359 (0.340)	0.357*** (0.00887)	0.705*** (0.0958)
P_{t-1}	255.0* (116.8)	-44.51 (292.3)	86.81** (27.66)	156.1* (74.45)
P_t	-734.6*** (132.0)	-731.0*** (136.5)	-243.0* (118.3)	-244.8* (107.6)
P_{t+1}	-42.20 (45.44)	-116.8 (72.95)	-13.54 (38.69)	20.21 (27.06)
Rational Addiction	No	No	No	No

Notes: The dependent variable in these models is present consumption. All models include individual level fixed effects, year and month factor variables, and controls for household income. Standard errors are clustered at the individual level. Instruments for the 2SLS estimator include P_{t-2} and P_{t+2}, following Dragone and Raggi(2021).In the AR1 model rational addiction is shown by a significant positive coefficient on C_{t-1} and a significant coefficient on P_{t+1}.

* p<0.05, ** p<0.01, *** p<0.001

Table 8. Estimates of AR(1) Models of Rational Addiction by BMI Group						
Regular Soda						
	OLS			2SLS		
	Normal	Overweight	Obese	Normal	Overweight	Obese
C_{t-1}	0.109***	0.0816***	0.210***	-0.0386	-0.797	0.409
	(0.0304)	(0.0110)	(0.0425)	(0.156)	(2.142)	(0.378)
P_{t-1}	825.5***	92.14	284.7**	863.4**	-333.4	484.0
	(204.6)	(141.4)	(94.35)	(262.8)	(1055.2)	(329.5)
P_t	-1016.8***	-546.3***	-986.3***	-1002.4***	-511.8*	-1060.3***
	(228.4)	(164.4)	(109.9)	(222.5)	(214.7)	(138.8)
P_{t+1}	-144.6	0.320	-9.446	-217.4	-152.3	52.46
	(175.5)	(49.59)	(87.64)	(190.9)	(342.3)	(128.3)
Rational Addiction	No	No	No	No	No	No
Low-Fat/Skim Milk						
	OLS			2SLS		
	Normal	Overweight	Obese	Normal	Overweight	Obese
C_{t-1}	0.326***	0.286***	0.344***	1.011***	1.097***	0.602***
	(0.0206)	(0.0103)	(0.0168)	(0.216)	(0.164)	(0.158)
P_{t-1}	23.15	-46.55	49.41	444.2**	515.9***	73.13*
	(58.94)	(41.19)	(26.89)	(169.9)	(130.7)	(36.37)
P_t	-592.5***	-581.0***	-136.4	-427.3***	-404.6***	-129.5
	(81.83)	(49.99)	(86.41)	(103.9)	(71.83)	(74.15)
P_{t+1}	-193.2**	-250.3***	5.349	8.126	25.40	13.63
	(59.76)	(39.30)	(12.50)	(124.1)	(82.66)	(18.00)
Rational Addiction	Yes	Yes	No	No	No	No

Notes: The dependent variable in these models is present consumption. All models include individual level fixed effects, year and month factor variables, and controls for household income. Standard errors are clustered at the household level. Instruments for the 2SLS estimator include Pt-2 and Pt+2, following Dragone and Raggi(2021).In the AR1 model rational addiction is shown by a significant positive coefficient on Ct-1 and a significant coefficient on Pt+1.

* p<0.05, ** p<0.01, *** p<0.001

Table 9. Average Results for Random Coefficients Estimation of AR(1) models of Rational Addiction

	Regular Soda	Low-fat/Skim Milk
C_{t-1}	0.097*** (0.004)	0.191*** (0.003)
P_{t-1}	1580.97*** (235.32)	848.55*** (102.99)
P_t	-3150.80*** (301.37)	-1102.80*** (122.63)
P_{t+1}	286.62 (223.70)	396.13*** (100.78)
Proportion of Households with Rational Addiction	1.01%	2.34%
Proportion of Normal BMI Households with Rational Addiction	0.97%	1.99%
Proportion of Overweight BMI Households with Rational Addiction	0.85%	2.67%
Proportion of Obese BMI Households with Rational Addiction	1.21%	2.21%

Notes: The dependent variable in these models is present consumption. Estimation procedures for parameters in this model are described in the methodology section. In the AR1 model rational addiction is shown by a significant positive coefficient on C_{t-1} and a significant coefficient on P_{t+1}.

* p<0.05, ** p<0.01, *** p<0.001