

# Behavioral Feedback: Do Individual Choices Influence Scientific Results?\*

Emily Oster, Brown University and NBER

December, 2018

## Abstract

In many health domains, we are concerned that observed links - for example, between “healthy” behaviors and good outcomes - are driven by selection into behavior. This paper considers the additional factor that these selection patterns may vary over time. When a particular health behavior becomes more recommended, the take-up of the behavior may be larger among people with other positive health behaviors. Such changes in selection would make it even more difficult to learn about causal effects. I formalize this idea in a simple model. I test for evidence of these patterns in the context of diet and vitamin supplementation. Using both microdata and evidence from published results I show that selection varies over time with recommendations about behavior and that estimates of the relationship between health outcomes and health behaviors vary over time in the same way. I show that adjustment for selection on observables is insufficient to address the bias. I suggest a possible robustness approach relying on assumptions about proportional selection of observed and unobserved variables.

## 1 Introduction

In many domains individuals face recommendations about behaviors - these include, for example, their financial choices or decisions about investments in children. This paper focuses on health behaviors, where it is common for individuals to face recommendations about how to improve their health (i.e. take vitamins, eat vegetables, exercise, and so on). The literature demonstrates

---

\*I am grateful for comments to Isaiah Andrews, Amy Finkelstein, Matthew Gentzkow, Ilyana Kuziemko, Matthew Notowidigdo, Jesse Shapiro, Andrei Shleifer, Heidi Williams and participants in a seminar at the US Census. I am grateful to Valeria Zurla, Marco Petterson, Claire Hug, Julian DeGeorgia, Sofia LaPorta, Cathy Yue Bai, James Okun and Geoffrey Kocks for outstanding research assistance. This paper benefits from the Nurses Health Study data, funded with NIH grant UM1 CA186107.

that adherence to these recommendations varies across individuals; for example, those with more education or income are more likely to engage in “positive” health behaviors (e.g. Berrigan et al, 2003; Friel, Newell and Kelleher, 2005; Finke and Huston, 2003; Kirkpatrick et al, 2012; Cutler and Lleras-Muney, 2010; Cutler, Lleras-Muney and Vogel, 2008; Goldman and Smith, 2002). One result is that it is a challenge to learn about these relationships in observational data, due to problems of omitted variable bias. This issue is well-known (e.g. Greenland et al, 1999; Vandenbroucke et al, 2007).

What is less frequently considered is the fact that the selection problem may change over time. In particular, in the presence of individual behavioral response to health advice, the degree of bias in estimates of these relationships may evolve with health recommendations.

To be concrete, consider a hypothetical case in which researchers are evaluating the relationship between pineapple and cardiovascular health and imagine that a particular study finds a small positive relationship between pineapple consumption and low cholesterol. In response to the advice, some people would increase their consumption of pineapple. It is plausible - even likely - that these would be the people who are most concerned about their health. But this group is also likely to be engaged in other cholesterol-lowering behavior. A result of this is that later studies of this relationship may see the bias in the estimates *exacerbated* by the increased selection.<sup>1</sup> Even if the initial finding is a statistical accident and the true effect is zero or even negative, later studies may find large positive effects.

These dynamics may be interesting for several reasons. First, it may be of *per se* interest to understand how behaviors change in response to changes in recommendations. This has been studied in some individual contexts, but not widely analyzed (e.g. Cutler, 2004; Chern et al, 1995; Brown and Schrader, 1990; Chang and Just, 2007; Roosen et al, 2009; Kinnucan et al, 1997; Ippolito and Mathios, 1995). Second, to the extent that these dynamics are important enough to change the estimated relationship between behaviors and outcomes, it suggests the need to consider variation in bias over time in studying these relationships. Further, it suggests that the process of research and publication may actually affect our ability to learn about relationships in data.

The primary goal of this paper is to explore whether the dynamics implied by the example occur in empirical settings. I do this using data on diet and vitamin supplementation.

I begin, in Section 2, by outlining a more formal version of the basic theoretical insight above.

---

<sup>1</sup>This discussion, and indeed this paper overall, presumes that the actual size of the causal effects is the same in each period. This seems reasonable since these are intended as biological relationships, unlikely to be changing within a population on a year to year time frame.

I focus on deriving conditions under which these dynamics would appear. A key result is that we expect to see these dynamics more strongly in behaviors which are less common in the baseline period. If a behavior is very rare, then there is room for more significant early adoption by individuals who engage in other health behaviors. This framework points to a number of signatures of these dynamics in the data, notably a correlation between changes over time in the selection of behaviors and changes over time in the relationship between behavior and health outcomes.

I then provide a number of pieces of evidence suggesting these dynamics are present in the data in the settings I consider.

I begin by showing changes in the selection of behaviors in the face of changing recommendations. I look at both vitamins and dietary patterns, using data from surveys (the NHANES) and from consumer scanner data (the Nielsen HomeScan panel). I show that when health behaviors (vitamin D and E supplements, sugar consumption, fat consumption, Mediterranean diet) are more recommended, they are more common among individuals who do other healthy behaviors (exercise, not smoking) and among individuals with higher education and income. Indeed, I show that the relationship between the behavior and these other health measures moves strongly together with the levels of the behavior.

I then turn to the relationship between behavior and disease and show that when the behaviors are more recommended, the links with health outcomes are also stronger. For vitamins, I show this result based on evidence from published work. For diet, I again use NHANES survey data to directly show the links in the microdata. This latter analysis allows me to overlay the selection relationship (behavior-confound) on top of the disease relationship (behavior-disease) and show they move together. This covariance holds even with sociodemographic controls included in the disease regressions. In this latter set of analyses I also show a similar co-movement of selection and the disease-behavior gradient for consumption of individual foods.

Finally, I use an additional *panel* dataset, the Nurses Health Study, to look in more detail at transition patterns around a particular behavior, vitamin E supplementation. I show first that the same patterns arise in these data, even though the demographic group is much more limited (i.e. only nurses). When supplementation is more recommended the relationship with not smoking is stronger. Further, the relationship with survival also strengthens; this bolsters the evidence from published work on vitamins and health outcomes. These data allow me to look at what drives this and find, not surprisingly, that the effects are driven by positively selected people transitioning into supplementation after it is recommended, and more negatively selected people transitioning out.

Overall, the graphical, regression and transition evidence shows clear evidence of the dynamics hypothesized in Section 2. This may be interesting for thinking about who responds to health recommendations, but it also suggests a strong note of caution in interpreting these effects and, in particular, in interpreting their variation over time. The act of research may, itself, be driving future research findings.

In the last section of the paper I consider whether it is possible to use the change in selection patterns directly to evaluate robustness and generate more plausibly causal estimates. In particular, I combine the analysis here with an assumption of proportional selection on observed and unobserved variables (Altonji et al, 2005; Oster, 2018). I show the multiple selection regimes may provide an opportunity to infer causal effects in this framework *without* some of the additional assumptions which that literature requires. I apply this approach to the analysis of dietary patterns; taken at face value, it points (for example) to robust impact of the Mediterranean diet, but less so for sugar.

This paper relates to literature within economics and elsewhere. Many authors have noted that observational evidence in health settings is often contradicted by randomized trials (Autier et al, 2014; Maki et al, 2014; Brownlee et al, 2010); this happens even in cases where a large observational literature exists prior to the trial.

There is further a large literature showing that more educated individuals and those of higher socioeconomic status are more likely to adhere to health recommendations, which is the motivation for the ideas here (e.g. Berrigan et al, 2003; Friel, Newell and Kelleher, 2005; Finke and Huston, 2003; Kirkpatrick et al, 2012; Cutler and Lleras-Muney, 2010; Cutler, Lleras-Muney and Vogel, 2008; Goldman and Smith, 2002). I also relate to a large literature on consumer responses to health information (e.g. Cutler, 2004; Chern et al, 1995; Brown and Schrader, 1990; Chang and Just, 2007; Roosen et al, 2009; Kinnucan et al, 1997; Ippolito and Mathios, 1995). Finally, the paper relates, although more tangentially, to the theoretical literature on fads and herding (e.g. Bikhchandani et al, 1992).

## 2 Theoretical Framework

In this section I briefly formalize the intuitive model of behavioral response described in the introduction. The goal here is to provide a model and some conditions under which these dynamics would arise; it is useful to note that other models of behavior may also produce such dynamics,

and there are certainly models under which we would not see them (and, indeed, parameters within this model where they do not appear). In that sense, this represents a possibility result designed to develop intuition. The first subsection describes a model of individual behavior. The second overlays on this a description of the process of research, and suggests implications for patterns that would arise in the data.

## 2.1 Model of Behavior

### 2.1.1 Setup

I consider a set of individuals who may undertake any of a set of health behaviors from a vector  $\Lambda = \{\Lambda_1, \dots, \Lambda_n\}$ . Assume each behavior is binary and defined such that a value of 1 indicates a positive health behavior. Health behavior  $j$  has a health value  $\kappa_j$ , and the overall health index  $h$  is a linear sum:

$$h = \kappa_1\Lambda_1 + \kappa_2\Lambda_2 + \dots + \kappa_n\Lambda_n.$$

The assumption of binary behaviors and a linear sum may seem restrictive, although we can imagine converting continuous behaviors to binary by discretizing them, and it would be possible to incorporate interactions between behaviors in a similar way.

Health behaviors are costly; we denote the cost of behavior  $x$  as  $c_x$  and assume it is the same across all individuals.

Individual  $i$  has concave utility over their health:  $U_i = U_i(h(\Lambda))$ . Individuals differ in their health valuation. We will define individual  $i$  as having a higher health value than individual  $j$  if  $U'_i(h) > U'_j(h)$  for all  $h$ .<sup>2</sup>

Each individual chooses their optimal value of  $h$ , trading off their utility value of health against the cost. For individual  $i$  we can write their problem as:

$$\max_{\Lambda} U_i(h(\Lambda)) - \sum_{k=1}^n c_k \Lambda_k$$

Consider now ranking behaviors  $r = 1, \dots, n$  based on  $\frac{c_r}{U_i(\sum_{j=1}^r \kappa_j \Lambda_j)}$ , which can be interpreted as the relative cost of undertaking behavior  $r$ . We will reference behaviors as  $\Lambda_{r=x}$ , with health value  $\kappa_{r=x}$  and cost as  $c_{r=x}$

---

<sup>2</sup>This is a slight abuse of notation since  $h$  depends on both discrete values ( $\Lambda$ ) and continuous values ( $\kappa$ ) so may in some cases not be defined.

Individuals will undertake the  $\Lambda_{r=1}$  behavior first - this is the behavior with the lowest cost per health unit. They will continue to add behaviors, moving up the  $r$  ranking, until the marginal behavior, which we denote  $\Lambda_{r=\tilde{r}}$ . Given that  $\Lambda$  is discrete, this marginal behavior is defined such that:

$$U_i \left( h = \sum_{j=1}^{\tilde{r}+1} \kappa_j \Lambda_j \right) - U_i \left( h = \sum_{j=1}^{\tilde{r}} \kappa_j \Lambda_j \right) < c_{\tilde{r}+1}$$

This is the last behavior for which the marginal cost is lower than or equal to the marginal gain.

Under this model, some individuals will undertake more health behaviors than others. Notably, those with higher health values will engage in more health behaviors. Define  $h_i^*$  as the health index corresponding to the optimal health behaviors chosen by individual  $i$ . Note we can write  $h_i^* = \sum_{j=1}^{\tilde{r}_i} \kappa_j \Lambda_j$ ; it will be convenient below to refer to this object. Note that a higher value of  $h_i^*$  corresponds to engaging in more health behaviors.

Individuals also realize some particular health outcomes - say, obesity or heart disease - which are a function of behaviors. Define a particular health outcome  $Y_i$  for individual  $i$ . We assume this outcome is a linear function of the  $\Lambda$  health behavior vector directly, so we write  $Y_i = \alpha + \vartheta \Lambda_i + \epsilon_i$ . Note that since elements of the  $\vartheta$  vector can be 0, it may be that some behaviors do not directly impact health outcomes.

### 2.1.2 Change in Value of Behavior

This paper is primarily concerned with the dynamics that occur when there is a change in the (perceived) value of a behavior. Consider a behavior  $A$  which previously had  $\kappa_A = \underline{\kappa}_A$ , indicating a low contribution to health. It is now announced that behavior  $A$  has a higher contribution to health:  $\kappa_A = \overline{\kappa}_A$ . Assume this occurs between a time  $t$  and  $t + 1$ . I will consider two changes between these periods. First, changes in the relationship between adoption of  $A$  and other behaviors. Second, changes in the relationship between  $A$  and health outcome  $Y$ .

I develop these results below.

**Behavior Selection Dynamics** I begin by considering adoption of behavior  $A$  after the change in recommendation. All proofs appear in Appendix B.

**Proposition 1** *Behavior  $A$  will be adopted by (weakly) more individuals following the change in perceived value.*

This first proposition is straightforward. Of more importance is the changes in selection in behavior  $A$ .

Recall the optimal health index from above,  $h_i^*$ . Because individuals have different health utility functions, this optimal index has a distribution in the population, and an average,  $E[h_i^*]$ . Denote  $h^A = \sum_{j=1}^A \kappa_j \Lambda_j$  as the threshold for adopting behavior  $A$ . This is not specific to any individual, since the costs and values  $\kappa_A$  are the same for all individuals. When  $h_i^* \geq h^A$ , individual  $i$  is engaging in behavior  $A$ .

**Assumption 1** *Assume:*

$$\frac{E_{t+1} [h^* | h^* \geq h_{t+1}^A] - E_{t+1} [h^*]}{E_t [h^* | h^* \geq h_t^A] - E_t [h^*]} \geq \frac{P(h^* > h_t^A)}{P(h^* > h_{t+1}^A)}$$

This indicates that the ratio of the difference in average health index between those who adopt behavior  $A$  and the average person, before and after the change in recommendation, is greater than the inverse of the ratio of the probability of adopting behavior  $A$  before and after the change in recommendation

**Proposition 2** *Under Assumption 1,  $Cov_{t+1}(A, h^*) \geq Cov_t(A, h^*)$ .*

The key result in this proposition is that the relationship between the behavior of interest -  $A$  - and other positive health behaviors will strengthen after the change in recommendation. Recall that  $h^*$  is the health index corresponding to the optimal choice, so higher values of this imply more health behaviors. Increases in the covariance imply  $A$  is newly adopted more frequently by those who undertake more health behaviors before the change.

Assumption (1) is key to this result. Some simulation intuition around this assumption is delivered in Appendix B. There are two central features to this intuition. First, behavior  $A$  must not be too common at baseline. If nearly everyone already engages in the behavior at baseline, further increases will diminish rather than exacerbate inequality. Second, the increase in value of  $A$  must not be too large, for effectively the same reason - if the result of the change is to cause everyone to adopt  $A$ , the covariance may not increase.

What this says is that we do not expect to see these dynamics in all settings. In that sense, it is an empirical question whether, in the settings I consider, we see these results.

Proposition 2 links behavior  $A$  to other health behaviors. In addition, we can consider the role of other covariates. Specifically, assume that we are able to observe a variable  $\omega$ , such that  $\omega$  is

correlated with  $U'$  and, as a result, with  $h^*$ .<sup>3</sup> This is intended to capture a variable like education or income.

**Proposition 3** *Under Assumption (1),  $Cov_{t+1}(A, \omega) > Cov_t(A, \omega)$ .*

**Disease-Behavior Dynamics** I turn now to the estimated relationship between behavior  $A$  and health outcomes.

**Assumption 2** *Assume that there is no treatment effect heterogeneity in the impact of  $A$  on  $Y$ .*

This is a strong assumption. I am interested in the possibility that changes in selection are driving changes in the estimated relationship between  $A$  and  $Y$  due to unobserved confounding. If the causal impact of  $A$  on  $Y$  is different for different groups, then the changing selection could change the relationship even in the absence of bias.

**Proposition 4** *Define  $\hat{\Lambda}$  as a strict subset of  $\Lambda$ . Under Assumptions (1) and (2) we can derive the following two results:*

(A)  $Cov_{t+1}(A, Y) > Cov_t(A, Y)$

(B)  $Cov_{t+1}(A, Y|\hat{\Lambda}) > Cov_t(A, Y|\hat{\Lambda})$ .

This says that as the behavior becomes more recommended, and thus the selection on the behavior changes, the effect we estimate of the behavior on health outcomes will change. This will be true even if we observe some of the confounding variables, as long as we do not observe all of them. Note if all elements of  $\Lambda$  were observed and controlled for then the estimated impact of  $A$  on  $Y$  with controls would not vary and we would estimate the true effect of  $A$  on  $Y$ .

A corollary relates this proposition to the behavioral selection dynamics.

**Corollary 1** *There will be a positive correlation between  $Cov(A, Y)$ ,  $Cov(A, \omega)$  and  $Cov(A, h^*)$ .*

This corollary says, simply, that the relationship between the behavior and outcome will move with the behavioral selection.

---

<sup>3</sup>This result follows from the monotonicity and concavity of  $U_i$ .



## 2.2 Research Process and Data Implications

The preceding subsection derives conditions under which we would see the dynamics described in the simple example in the introduction. Here, I consider overlaying a research process over these results, and ask how we can look for these patterns in the data.

I consider the research process evaluating the effect of behavior  $A$  on outcome  $Y$ . It will be helpful to pull out the particular behavior  $A$  in the estimating equation for  $Y$  and specify the true data generating process as

$$Y_i = \alpha + \beta A_i + \vartheta \mathring{\Lambda}_i + \epsilon_i$$

where  $\vartheta$  and  $\mathring{\Lambda}_i$  represent the coefficient and behavior vectors with behavior  $A$  removed.

Assume the research process is as follows. In each year, researchers draw a sample of individuals and collect data on behavior  $A$ , outcome  $Y$  and a set of other variables  $\Phi$ . This vector  $\Phi$  may include some elements of  $\mathring{\Lambda}$ , along with elements of  $\omega$  (other demographics, etc). Following the data collection, they estimate the effect of  $A$  on  $Y$  using the feasible equation below, where we have introduced subscripts  $t$  to indicate the estimation is specific to a year. Note that  $\beta$  also has a  $t$  subscript to indicate that it may not be equal to the true  $\beta$ .

$$Y_{it} = \alpha_t + \beta_t A_{it} + \varsigma \Phi_{it} + \epsilon_{it}$$

Typically, researchers then report  $\beta_t$ .

In the context of this setup, I consider cases where information about the health value of  $A$  changes and look for signatures of the above dynamics in the data.

The first set of results focuses on the selection patterns directly. Propositions 2 and 3 highlight the dynamics of the relationship between  $A$ , other health behaviors and elements of  $\omega$ . I use the data to look directly for these changes in selection patterns which would be implied by Assumption (1). I will often refer to the relationship between  $A$  and other health behaviors or elements of  $\omega$  as the “behavior-selection gradient”, understanding that this is not a technical use of the term “gradient”.

In the second set of results I focus on the time-varying relationship between behavior  $A$  and outcome  $Y$ . As detailed in Proposition 4, if we cannot observe all elements of  $\Lambda$ , the estimated  $\beta_t$  effects will vary over time and, as in Corollary (1) will move with the behavior-selection gradient. I will refer to the relationship between health and behavior as the “health-behavior gradient”, with a similar note on the terminology.

It is important to note that if we observe all the elements of  $\Lambda$ , or if the demographic controls are sufficient to fully explain  $\Lambda$ , then these latter dynamics will not occur. In that case, we expect the  $\beta_t$  coefficients to be the same in each period, and equal to  $\beta$ . In this sense, observing that they are different provides an (indirect) test for whether the included controls fully capture the omitted factors.

A final note is that Corollary (1) suggests we may look for co-movement in the two gradients directly.

### 3 Data

There are four key data elements for this paper. First, a target health behavior for which we are interested in identifying the effect (denoted by  $A$  above). Second, an additional set of behaviors which may also influence the health outcome ( $\mathring{\Lambda}$  above). Third, covariates which influence the taste for health ( $\omega$  from above). Finally, a set of health outcomes possibly linked to this behavior ( $Y$ ). Some of the patterns detailed above can be explored even with only a subset of these data.

Below, I detail the key variables and the datasets used.

#### 3.1 Key Variables

**Target Health Behaviors: Vitamins** I look at two vitamin supplements: Vitamin D and Vitamin E.

**Target Health Behaviors: Diet** The primary analysis looks at three components of diet: sugar as a share of carbohydrates, saturated fat as a share of total fat and an index of adherence to a Mediterranean diet. In addition, in a secondary analysis I will focus on estimating the impacts of individual foods.

**Other Health Behaviors** The primary other health behavior I focus on is exercise. In addition, I look at smoking and, in the case of vitamins, a metric of overall diet quality.<sup>4</sup>

**Proxies for Health Value** I use socioeconomic status - the first principal component of education and household income - as the proxy for overall health value. This is in line with the literature, which tends to include these variables as controls in the analysis.

---

<sup>4</sup>I do not look at vitamin-taking behavior as another metric when I consider diet, since vitamins are only sometimes recommended during this period.

**Health Outcomes** For vitamins, the health outcome considered is cancer. For diet, I look at cardiovascular health, BMI and obesity.

## 3.2 Data Sources

Below I describe the data sources in detail. In addition, Appendix Table A1 summarizes the datasets used and the parts of the analysis for which they are relevant.

### 3.2.1 NHANES

The National Health and Nutrition Examination Survey (NHANES) is a nationally representative survey which has been run, in some form, since the 1960s. In this project, I use data from the NHANES III (1988 through 1994) and from the continuous NHANES (beginning in 1999/2000 through 2012/2013).

Information on vitamin supplementation is obtained from the vitamin supplement modules. I focus on individual vitamin supplements - that is, is someone taking a single-ingredient Vitamin D or E supplement. Information on diet is generated from the daily dietary recalls in the study. In the case of the Mediterranean diet I generate a Mediterranean diet score as described in Trichopoulos et al (2003).<sup>5</sup>

I extract data on education, income and other demographics from the demographic survey portion of the NHANES. The NHANES also provides a measure of vigorous exercise, which I standardize within year, as well as information on smoking behavior.

To study health outcomes, I extract information on cardiovascular health, BMI and obesity from the data. I construct an index of heart health based on blood pressure and cholesterol. A significant advantage of the NHANES data is that all of the health measures are collected objectively - from weighing and measuring the individual and testing their blood pressure, etc. This avoids issues of recall bias. We will not be able to study cancer development as an outcome here, so our analysis of vitamins in the NHANES will be limited to analyzing selection patterns in vitamin taking.

### 3.2.2 Nielsen HomeScan Data

The Nielsen HomeScan panel tracks consumer purchases using at-home scanner technology. Households that are part of the panel are asked to scan their purchases after all shopping trips. The

---

<sup>5</sup>The score assigns a value of 0 or 1 in nine dietary elements, where a value of 1 is given if someone is either above the median in a good food category (vegetables, fish, etc) or below the median in a detrimental food category (dairy, meat, sugar).

Nielsen data records the UPC of items purchased. Einav, Leibtag and Nevo (2010) validate the reliability of the HomeScan panel. I use Nielsen data available through the Kilts Center at the University of Chicago Booth School of Business. These data span 2004 through 2016.

These data will be used to look at selection in vitamin purchases over time. They do not contain information on health, and variation in nutrient data coverage over time makes it difficult to analyze diet. However, it is possible to look at vitamins, in particular by generating variables indicating whether the household purchased each vitamin supplement during each year. Information on household education and income can then be used to analyze selection. In addition, I incorporate data on cigarette purchases as a metric of other health behaviors.

### **3.2.3 Publications**

In the analysis of vitamins I also draw information from published work on the relationship between vitamin supplements and cancer. I locate publications in two ways. First, I scrape Pubmed for “Vitamin X and cancer” and extract relevant studies, limiting to studies in journals in the top 20% in terms of impact factor. Second, I extract lists of publications from meta-analyses of these relationships. The former of these ensures I do not miss important studies which have not been included in meta-analyses. Most of the citations, however, come from the meta-analyses, where the original authors have carefully extracted all relevant studies. I focus on observational studies and exclude RCTs.

For each original study I then extract information on the treatment (either vitamin D or vitamin E supplementation), the outcome (a type of cancer), the years of data covered in the study, the population characteristics and, importantly, whether the study findings were significant. I focus on significance rather than magnitude because given the varying approaches across studies, and the varying types of cancer, it is difficult to compare magnitudes.

In all, the resulting dataset includes 82 studies of vitamin D supplementation, and 83 studies of vitamin E supplementation.

### **3.2.4 Nurses Health Study**

In Section 6 I use data from a panel of nurses, the Nurses Health Study (NHS), to analyze selection patterns in Vitamin E, in particular, and to look at transitions into and out of Vitamin E supplementation around the recommendations. The NHS recruited a cohort of approximately 120,000 female nurses in 1976 and conducted by-mail surveys of the cohort every two years. The

study is ongoing and response rates are very good. Information is available on dietary patterns, some behavioral characteristics and outcomes, including mortality. The cohort is described in more detail in Colditz et al, 1997.

I use data from 1982 to 2006 on vitamin E supplementation, smoking behavior and death rates. These will supplement the primary analysis by allowing me to look at characteristics of those who transition in and out of supplementation around recommendations. In addition, linking Vitamin E to mortality will allow me to supplement the evidence from publications.

## 4 Results: Selection into Behaviors

In this first section of results, I look for changes in the selection into behaviors over time for both vitamins and diet. Referring back to the theory, this section is focused on looking for patterns echoing Propositions (1)-(3).

### 4.1 Vitamins

To begin the analysis, I identify timing of significant information events for each vitamin supplement. These events include changes in government recommendations, the advice of national organizations and major research findings. Data from Google Trends are used to validate these events where possible. Appendix Table A2 describes these events in detail. Although the discussion in Section 2 is agnostic about the source of changes in popularity of the behavior, these events provide an anchor for understanding patterns in the data.

I look for evidence of co-movement between the levels of vitamin supplementation and the selection over time. I focus on estimating selection as either (a) the relationship between supplementation and exercise or (b) the relationship between supplementation and socioeconomic status (the first principal component of education and income). I regress health behaviors on this socioeconomic status variable in each year of the data (these regressions include controls for age and gender). I then look at the relationship between these coefficients and the levels of the variables over time.

Graphical analysis of these results are shown in Figure 1. The first four graphs use data from the NHANES; the bottom two use data from HomeScan. Note that I do not observe exercise in the HomeScan data, so look only at socioeconomic status gradients there.

The graphs show strong evidence of co-movement between levels of behavior and either gradient.

For example, in the case of vitamin D there are increases in the gradient as the behavior got more popular (after some positive research findings) and then a subsequent decrease as the popularity waned after research findings suggesting the benefits were overstated.

We can turn also to regression evidence which parallels the graphs in Figure 1 but incorporates other gradient variables in addition to exercise and socioeconomic status. To do this, I aggregate across the vitamin measures, and create a dataset in which each observation is a vitamin supplementation behavior-year and the dataset includes the level of the behavior (i.e. the share of people who take the supplement), and the gradient with respect to exercise, socioeconomic status, smoking and diet quality (i.e. the coefficient from a regression of taking the supplement on each of these variables). I then run regressions of these gradients on behavior levels, including fixed effects for each vitamin.

The results are shown in Table 1. Each cell represents a different regression, with Panel A focusing on the NHANES and Panel B focusing on the HomeScan. The relationships between levels and gradients are positive and significant in all cases. In other words, the data show that as these behaviors become more popular there is a stronger positive link between the behavior and other health measures.

## 4.2 Dietary Patterns

I analyze the three dietary patterns - sugar intake, fat intake and Mediterranean dietary patterns - in the same way as vitamins. Appendix Table A2 describes salient information events in detail.

Graphical analysis of these dietary results appear in Figure 2. The top panels relate the behavior levels to the gradients with respect to exercise, and the bottom with gradients with respect to socioeconomic status. As in the vitamin case, we see clear co-movement here. This is most notable in the case of sugar and Mediterranean diet, where in the most recent years there is a large change in behavior which corresponds to a large movement in the estimated gradient.

Table 2 shows the regression evidence corresponding to these figures, including the smoking gradient. Again, we see significant relationships. As with vitamins, there is a stronger positive link between behavior and other health measures as the behavior becomes more recommended.

## 5 Results: Selection Patterns and Behavior-Outcome Relationships

I turn now to variation in the relationship between these health behaviors and health outcomes. I begin again with vitamins, where the evidence is drawn from published results, and then move on to my own analysis of micro-data in the case of diet.

Broadly, the approach in this section is to explore whether the estimated impact of behavior on outcomes (the “behavior-outcome gradient”) moves in tandem with the relationship between the selection patterns (the “selection gradient”).

### 5.1 Vitamins

The key health outcome associated with vitamin supplementation is the development of cancer. The behavior-outcome gradient of interest, then, is the relationship between vitamin supplementation behavior and subsequent cancer incidence. I explore changes in this behavior-outcome gradient over time using data from published work on the link between vitamins and cancer. On the one hand, this is less direct than looking at these relationships in the microdata (as I will do below for diet), since the publication process introduces additional noise into the data. On the other hand, this is a very direct test of whether these dynamics affect the process of scientific discovery.<sup>6</sup>

This analysis uses the data described in Section 3. For each publication, I allocate the study to time periods based on the timing of the data in the study. Some studies are allocated in parts to different time periods. For example, if a study includes data from 2005 through 2010, I assign a 50% weight to the period before 2008, and a 50% weight to the period after. I then summarize the (weighted) share of significant cancer reductions across time periods. I residualized results with respect to the type of cancer studied (this does not affect the findings).

Note that the coding relates to the date of the data *not* the date of publication. Publication timing may be independently interesting, especially if we think that publications reflect conventional wisdom, but the focus here is on the data timing. Further, publications lag data, and as a result it is not feasible to look at results from the most recent time periods. In particular, in the case of Vitamin D we do not have data past 2010, and in the case of Vitamin E, we do not have sufficient data past 2004.

---

<sup>6</sup>In Section 6, when I turn to the Nurse Health Study, I will show using microdata that the links between Vitamin E and death rates overall do track the changes in supplementation patterns. This echoes and reinforces these results.

The results are shown in Figure 3, and are consistent with the changes in information outlined above. For Vitamin D, published results with data post-2007 are much more likely to find significant relationships between Vitamin D supplementation and cancer. For Vitamin E, a similar pattern occurs for results with data post-1993.

These line up with the changes in selection. The post-2007 period for Vitamin D is the period with the increase in the selection gradient; since the data ends in 2010, this does not coincide with the popularity decrease. Similarly, for Vitamin E we are seeing here the period from 1993 to 2004 when we believe the popularity of Vitamin E was higher and the selection gradient increased. The post-2004 period is not covered.

This evidence is suggestive of the patterns described in Section 2. Importantly, it suggests that these patterns appear in published work, where authors have attempted (in most cases) to control for some of the confounding behaviors and where the publication process has vetted the work. I turn now to the case of diet, where it is possible to conduct a direct analysis of these patterns in the microdata.

## 5.2 Dietary Patterns

The estimated impacts of dietary patterns on outcomes in the NHANES data vary over time. This can be seen in Appendix Table A3 which shows the range of estimates of the impact of behavior on various health outcomes across years in the NHANES. The range of impacts is quite large and, in a number of cases, includes both positive and negative values. This variation is not just noise, as we can typically reject equality of coefficients across years.

The existence of variation over time does not directly point to a relationship with selection. To test this directly, Figure 4 graphs the BMI gradient (regression of BMI on behavior) over time, alongside the exercise and socioeconomic status gradient over time. We redefine BMI as *negative* BMI so the theory predicts the series will go up and down together. Although these are noisier than the figures which focus on gradients and levels, we still see the series moving together. The link appears with both exercise and socioeconomic status gradients.

Table 3 shows regression evidence on these co-movements, parallel to Table 2. Here, I look at three outcome gradients, and three proxy gradients. For example, to create the top left cell I first regress BMI on each diet measure in each year, and extract the coefficients. I then regress diet on socioeconomic status in each year, extract these coefficients and match them to the BMI-diet coefficients at the yearly level. This resulting dataset has a behavior-year unit of observation.



Using these data I finally regress the behavior-outcome gradient on the behavior-socioeconomic status gradient, controlling for behavior fixed effects. I repeat this procedure for varying outcomes, and for socioeconomic status, exercise and smoking.

To create Panel A of Table 3 I run the disease regressions with no controls other than age and gender. This effectively tests claim (A) in Proposition 4 that the unconditional relationship between  $A$  and  $Y$  is increased after the change in recommendation. In 8 of 9 cases the gradients move significantly together. For example, in that first cell the coefficient is negative and highly significant, showing that when the behavior is more positively selected on exercise behavior, the link between the behavior and BMI is more negative. In the case of smoking we see the links with BMI and obesity, although the link with heart health is in the expected direction but not significant.

Panel B of Table 3 shows the same relationships, but in this case the regressions estimating the disease gradients include controls. This parallels claim (B) in Proposition 4. In these regressions I include standard demographic controls which are in the literature (e.g. Loftfield et al, 2018; Yang et al, 2014; Bao et al, 2013). If these controls are sufficient to address most or all of the bias then these gradients should be eliminated by their inclusion.

The estimated gradients are smaller in Panel B than Panel A, but they are still sizable and significant in all but one case. This indicates - consistent with the publication results - that these standard controls are insufficient to fully address the changes in selection which are occurring.

### 5.3 Individual Foods

As a final piece of evidence in this section, I consider a less event-driven approach to the data. Specifically, recall Corollary (1), which indicates that these dynamics imply the co-movement of the behavior-selection gradient and behavior-outcome gradient. We expect to see this even without taking a stand on the particular events that drive the selection gradient changes. Put differently: the analysis above relies, at least in part, on identifying particular events that drive changes in selection. But we can also look directly for changes in the behavior-outcome gradient, and ask whether they line up with changes in the selection gradient.

I do this by looking at consumption of individual foods in the NHANES. I mine the data for the subset of foods that show significant changes in their effects on BMI over time. I define four time periods, based on roughly dividing the data into quarters: 1988-1991, 1999-2003, 2005-2009 and 2011-2013. I then select for the sample of individual foods which either (a) shows significant differences in effects in at least 3 of 6 pairwise differences over time or (b) shows a significant linear

trend effect over time.

This approach identifies 33 individual foods (from an original set of 188). Selected foods include foods at the basis of the Mediterranean diet like nuts, seafood, seeds and legumes, high-sugar foods like soda, but also foods like margarine, peanut butter, berries and white and brown rice. Appendix C lists the full set of foods identified and the estimated impact of each one over time.

Using these foods, I estimate the co-movement between the selection gradient and the behavior-outcome gradient. To do this, I echo the analysis above, and estimate the behavior-outcome gradient for each food-time period, alongside selection gradients (with respect to socioeconomic status, exercise and smoking).<sup>7</sup> The resulting dataset is at the food-time period level, and I can estimate the relationships between the gradients at this level, including controls for the particular foods.

I summarize the results in Figure 6 and Table 4. Note that to construct these figures the behavior-outcome regressions do not include demographic controls, but Appendix Figure A1 and Table A4 show the same results with controls in those regressions.

Figure 6 shows a visual of the relationship between the gradients, with each set residualized with respect to the individual food effects. In this case, the health outcome visualized is BMI and the selection gradient is with respect to socioeconomic status. There is a strong downward sloping relationship visible in the graph. When the relationship between each food consumption and socioeconomic status is stronger, the estimated relationship between the foods and BMI is greater.

Table 4 shows the regression equivalent of this figure, echoing Table 3, and incorporating the other health outcomes and selection gradients. Again, we can see similar patterns to the results above. The gradients move significantly together in many cases; the effects are particularly robust when we consider smoking behavior.

The story for why we see these changes is less clear than in the first two analyses. In those cases we directly identified information events. Here, there are some foods - nuts, for example - where it seems clear what the information event is (namely, the same Mediterranean diet studies cited above) but in others - white rice, for example - it is less obvious. Nevertheless, the underlying point about dynamics survives the move away from these information events.

---

<sup>7</sup>In this case, the selection gradients are calculated based on a probit model, since some of these foods are (very) infrequently purchased, leading to misleading behavior in the linear probability model.

## 6 Nurse Health Study: Evidence on Transitions

I turn now to a final piece of evidence on these dynamics, using data from the Nurses Health Study on Vitamin E consumption. As noted above, the Nurse Health Study is a panel dataset which has followed a cohort of nurses over a period of more than 30 years, collecting information on their health and health behaviors. This cohort is, in some ways, less well suited to look at these issues than the NHANES. Most notably, this study contains only nurses, which constrains the range of socioeconomic status. In addition, the data does not contain information on family or personal income, so it is not possible to look for variation along the dimensions. Finally, many of the changes in recommendations exploited above (for example, the sugar and Mediterranean diet changes) occur when this cohort is very old and many of them have died.

However, a significant advantage of these data is the panel format, which allows me to look directly at *changes* in behavior and how these changes vary across groups. I do this in the context of Vitamin E supplementation, where the changes in recommendations occurred in the 1990s and early 2000s, when this cohort was still largely intact.

I begin by showing that even within this fairly homogenous cohort, we observe variation in selection which lines up with variation in behavior. This can be seen in Figure 5. In each sub-figure I show the share of the population who takes Vitamin E supplements alongside either the smoking gradient or the survival gradient. In both cases I define the outcomes to be positive: “not smoking” and “survival”, so the theory and earlier results suggest the series should move together. In the latter case, the outcome is survival for the 10 years following the survey. All gradients control for age fixed effects (but not for gender, since the whole cohort is female).

There are clear trends in Vitamin E supplementation, which align with the events in Appendix Table A2. Between 1992 and 1994, after the 1993 release of the pro-Vitamin E studies, the share of the population who takes Vitamin E begins to climb, ultimately increasing from about 20% to over half. It is notable that this is much higher than the general NHANES population, likely reflecting the fact that the sample is already selected. After 2004, when there was bad news about Vitamin E, the share taking supplements drops significantly.

These trends in supplementation rates are largely tracked by the gradients. This is most striking in the survival figure; the gradients track very closely on both the increase and then the decline in 2006. In the periods when Vitamin E is more popular, it appears to provide more survival benefits in this population. This reinforces the publication results on Vitamin E above.

For smoking we see again similar patterns around the two events, although the smoking gradient is elevated as well in the earliest period.

Thus far, this echoes the earlier results. In addition, we can look in more detail at what drives these changes. Panel A of Table 5 shows the smoking rates for four groups around the period of Vitamin E adoption. In particular, I define a “before” period (1990 to 1992) and an “after” period (1998 and 2000) and create a two by two table. The upper left cell shows the share smokers in the earlier period among people who take Vitamin E both before and after the change in recommendation, the bottom right shows those who never take it, and the two off-diagonal cells shows the changers.

Those who always take vitamin supplementation are much less likely to smoke in the baseline period than those who never take supplements (12.9% versus 18.7%). Two things drive the change in the gradient. First, the new adopters between the early and later period have lower smoking rates (13.7% versus 18.7%) than those who do not adopt (i.e. are “No” in both periods). Further, those who dis-adopt (stop taking the vitamin) after the positive recommendation are more likely to be smokers than those who keep taking it (14.1% versus 12.7%). Together, these dynamics drive the change in gradient.

Panel B of Table 5 shows death rates for the same groups. In this case, the data is limited to those who were alive in 1998 or 2000, and the outcome is death rates over the ten years following that. We see similar patterns here. For example, among those who do not use Vitamin E in 1990-1992, the death rate is lower (10.1% versus 10.4%) among those who switch to using after the change in recommendations. Similarly, the death rate for new adopters is lower than the death rate for dis-adopters (10.1% versus 11.5%).

Overall, these results supplement the above. They illustrate the patterns of behavioral adoption which must underlie the broader evidence we see in Section 5. Further, the evidence on vitamin E and death rates corroborates the publication data.

I return now, in the next section, to the NHANES data and the question of whether we can make any progress on robustness using the changes in selection.

## 7 Selection on Unobservables Approach

The paper, thus far, makes two central points. First, I document clear changes in the selection of behaviors over time, as they become more or less recommended. Second, I show evidence suggest-

ing these changes in selection change the estimated - and reported - impacts of these behaviors. Demographic controls are not sufficient to address these issues, and the publication process does not appear to fully constrain them.

This set of results should, first and foremost, add to our caution in interpreting observational results in settings like this. The problem of omitted variable bias is well known, but these results suggest such bias may be dynamic and, indeed, may *respond* to research findings. This suggests that awareness of the changes in recommendations over time should inform discussions about the plausible degree of bias in estimates.

In this section I consider the possibility that this dynamic selection also presents an opportunity. Specifically, these changes in selection could be used, along with a strong assumption about the selection on observed versus unobserved variables, to obtain better estimates of causal effects. This may be an approach to robustness in situations where selection varies over time. Below, I briefly formalize this framework and then show an example of these calculations in the context of dietary patterns.

## 7.1 Framework

I return to the framework in Section 2. I modify slightly by defining  $\bar{\Lambda}_i = \vartheta \hat{\Lambda}_i + \epsilon_i$  and rewriting the full equation for  $Y$  as

$$Y_i = \alpha + \beta A_i + \bar{\Lambda}_i$$

Effectively this means  $\bar{\Lambda}$  is a factor equal to the full vector of behaviors multiplied by their true coefficients, and the noise term. We can think of this as assuming the noise term is a “health behavior”, or assuming there is no true “noise” in health determination.<sup>8</sup>

We assume now that Assumption (1) holds and, therefore, between two periods we have an increase in  $Cov(A, h^*)$ .

In Section 2 I described  $\Phi$  - the control set - as containing some elements of  $\Lambda$  and  $\omega$ . I now explicitly define  $\tilde{\Lambda}$  as the vector of omitted factors, namely the residual from a regression of  $\bar{\Lambda}$  on  $\Phi$ . We can rewrite the true equation as

$$Y_i = \alpha + \beta A_i + \Pi \Phi_i + \tilde{\Lambda}$$

---

<sup>8</sup>The reason to redefine in this way is that when I move to the proportional selection adjustment below, this eliminates a free parameter (the maximum R-squared in the full regression). Given the estimation approach here, this is without loss of generality since this parameter is jointly determined with the degree of proportionality. This approach echos the setup in Altonji et al (2005).

where  $\Pi$  is the coefficient on  $\Phi$  in a regression of  $\bar{\Lambda}$  on  $\Phi$ . The feasible estimating equation is

$$Y_i = \alpha + \hat{\beta}A_i + \hat{\Pi}\Phi_i + \varepsilon_i$$

It is straightforward to see, by standard omitted variable bias logic, that  $\hat{\beta}$  is (possibly) biased and, in particular, that  $\hat{\beta} = \beta + \frac{Cov(A_i, \tilde{\Lambda}_i)}{Var(A_i)}$ .

Thus far this reiterates what I derived in Section 2. Now, however, I consider introducing an assumption from Altonji et al (2005) and Oster (2018) - namely, that there is proportional selection between the observed and unobserved variables. The formal assumptions is

$$\delta \frac{Cov(\Pi\Phi, A)}{Var(\Pi\Phi)} = \frac{Cov(\tilde{\Lambda}, A)}{Var(\tilde{\Lambda})}$$

This describes a relationship between the selection on the observed factors ( $\Pi\Phi$ ) and the unobserved factor ( $\tilde{\Lambda}$ ). As noted in the original work, this is a very strong assumption. In this context, the assumption would be delivered (for example) if the set of behaviors we observe is a random subset of all of the behaviors. Intuitively, work which uses this assumption tends to defend it by arguing the observed controls are a noisy proxy for the unobserved ones, which connects this directly to other formulations which rely more directly on this noise (e.g. Pei, Pischke and Schwandt, 2018).

With this assumption, the existing work shows that  $\beta$  is a function of  $\delta$  and a vector  $\mathbf{X}$  of parameters from the data (including controlled and uncontrolled regression coefficients and R-squared values).<sup>9</sup> The existing work imagines a single estimation, and that authors will combine the observed elements of  $\mathbf{X}$  with an assumption about  $\delta$  to comment on robustness of  $\beta$ .

In this case, however, we have observed multiple  $\mathbf{X}$  vectors, under varying selection regimes. With an assumption of constant  $\beta$  and  $\delta$  over time, we can write:

$$\beta = f(\delta, \mathbf{X}_1)$$

$$\beta = f(\delta, \mathbf{X}_2)$$

This is an identified system of equations. Effectively, this says we can infer the true  $\beta$  from the data by using the changes in selection. This is the exercise we will do below. Note that with more

---

<sup>9</sup>In Oster (2018) the results are developed under the assumption of noise in the data generating equation for  $Y_i$  and therefore contain an additional calibration parameter,  $R_{max}$ , which is the hypothetical R-squared in the true regression. We abstract away from that here by assuming no error. In practice,  $\delta$  and  $R_{max}$  are not individually identified and separating them is useful since it is easier to develop intuition about their values, but for the purposes here that is not relevant.

than two periods, the system is over-identified (since there will still be two unknowns -  $\beta$  and  $\delta$  - but more than two equations).

There are a large number of caveats here. This relies on the proportional selection assumption, which may not hold. It also relies crucially on the assumption that  $\beta$  and  $\delta$  are constant over time and that there is no treatment effect heterogeneity in  $\beta$ . The results above also rely on this last assumption, but since they are largely descriptive, this reliance is perhaps less important. Given that these relationships are biological, the assumption of treatment effect homogeneity may not be as much of a stretch as in some types of policy analysis.

Putting these together, it should be clear we see this as an exploratory or robustness exercise, rather than an alternative approach to inferring causality. This could be useful, for example, in meta-analytic approaches which rely on data from multiple time periods with differential selection. This approach adds to other approaches which incorporate additional information from the data to discipline this selection on unobservables approach to robustness (e.g. Finkelstein, Gentzkow and Williams, 2018).

## 7.2 Implementation and Results

To implement, I focus on the three dietary patterns in the data where I undertook the fullest analysis - sugar, fat and the Mediterranean diet. Broadly, the approach here is to ask what values of  $\delta$  and, more importantly,  $\beta$ , are most consistent with the observed data.

For each treatment I aggregate individual years into two (fat and Mediterranean diet) or three (sugar) periods.<sup>10</sup> It is possible, of course, to run this analysis for all years, but there are some challenges with sampling variability. By aggregating I increase the precision of this analysis but maintain the differences across years in selection. I use a grid search to identify the values of  $\beta$  and  $\delta$  which generate the smallest variation across periods in estimates of the causal effect after the proportional selection adjustment. In the cases with two periods, this is exactly identified. In the case of sugar, where there are three, it is over-identified. The non-linearity in the calculations and the over-identification mean that there may not be a single set of values that are the solution. I search for the smallest range of  $\delta$  values which deliver a single value of  $\beta$ .

A key input is the set of observables - i.e. what should be included in  $\Phi$ ? Here, I include only controls for socioeconomic status (education, income, race, marital status). This is done, as in the earlier disease regressions, to better echo the existing public health literature. Age and gender are

---

<sup>10</sup>I divide into periods based on a visual sense of the number of different “regimes” in the data.

adjusted for in all regressions, including the ones without the additional controls.

Table 6 reports the results for each health/behavior pair. The first columns report the coefficients in each of the sub-periods of the data. There is substantial variation over time. The fourth column reports the “best” causal estimate after adjustment for selection on unobservables, meaning the value that minimizes the different estimates of  $\delta$  across periods. I derive significance from a bootstrap. Finally, the last column reports the  $\delta$  value (or the range). This  $\delta$  does not have a clear interpretation, especially given that the error term is included in the unobservables. The values are generally small, which reflects this error conclusion.

The results here suggest some significant effects, although they are more limited in many cases than would be implied by the most optimistic yearly estimates. Increases in sugar consumption do appear to reduce heart health measures, but do not impact obesity or BMI. The Mediterranean diet has the most consistent effects after adjustment. This table also makes clear some of the less ideal features of this approach - for example, in a couple of cases the implied effects are very different than any of the controlled effects, which is probably not realistic. Nevertheless, it is perhaps comforting that the Mediterranean diet - the only diet pattern among these with support in randomized controlled trials - is the most consistent after the adjustment.

Again, given the strength of the assumptions required here I will stop well short of suggesting this as an primary approach to causality. However, it may prove to be a useful robustness tool.

## 8 Conclusion

In this paper I analyze the role of health behavior change in driving the selection features of health behaviors. I outline a simple data generating process in which changes in health recommendations differentially change health behaviors for different groups and show that these changes may influence estimated relationships between behavior and health over time. Using data on vitamin supplementation and diet I demonstrate that these dynamics occur in data. The degree of selection in behaviors varies over time, and the relationship between behavior and health also varies with these changes in selection. Finally, I suggest that using the changes in selection along with an assumption of selection on unobservables may be a useful approach to robustness.

This paper focuses on health behaviors and health outcomes, but the dynamics here may be present in other settings (parental behaviors, for example) where individual choices vary over time. The approach here may apply in those setting as well.



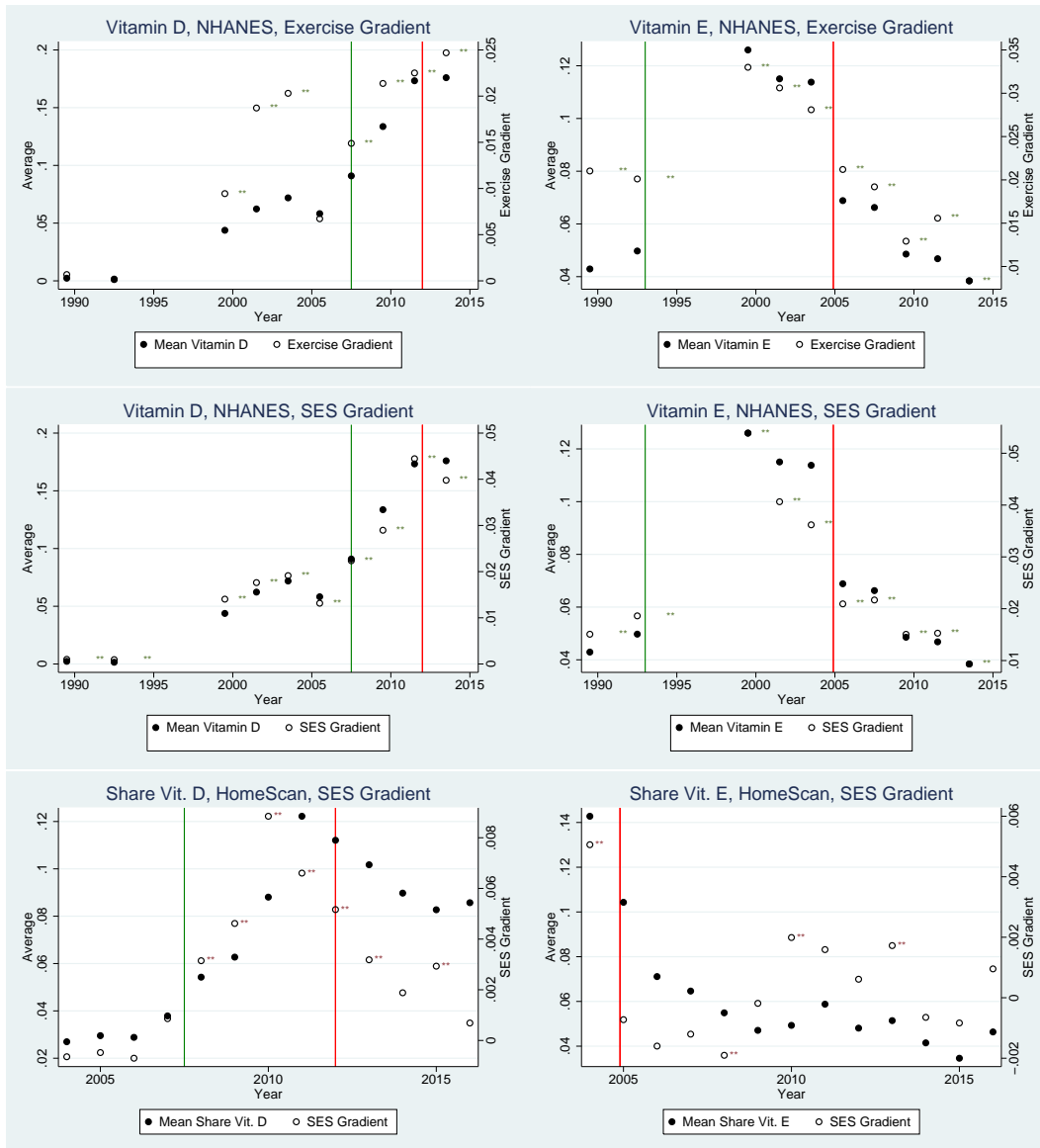
## References

- Altonji, Joseph G, Todd E Elder, and Christopher R Taber**, “Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools,” *Journal of political economy*, 2005, *113* (1), 151–184.
- Autier, Philippe, Mathieu Boniol, Cécile Pizot, and Patrick Mullie**, “Vitamin D status and ill health: a systematic review,” *The lancet Diabetes & endocrinology*, 2014, *2* (1), 76–89.
- Bao, Ying, Jiali Han, Frank B Hu, Edward L Giovannucci, Meir J Stampfer, Walter C Willett, and Charles S Fuchs**, “Association of nut consumption with total and cause-specific mortality,” *New England Journal of Medicine*, 2013, *369* (21), 2001–2011.
- Berrigan, David, Kevin Dodd, Richard P Troiano, Susan M Krebs-Smith, and Rachel Ballard Barbash**, “Patterns of health behavior in US adults,” *Preventive medicine*, 2003, *36* (5), 615–623.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch**, “A theory of fads, fashion, custom, and cultural change as informational cascades,” *Journal of political Economy*, 1992, *100* (5), 992–1026.
- Brown, Deborah J and Lee F Schrader**, “Cholesterol information and shell egg consumption,” *American Journal of Agricultural Economics*, 1990, *72* (3), 548–555.
- Brownlee, Iain A, Carmel Moore, Mark Chatfield, David P Richardson, Peter Ashby, Sharron A Kuznesof, Susan A Jebb, and Chris J Seal**, “Markers of cardiovascular risk are not changed by increased whole-grain intake: the WHOLEheart study, a randomised, controlled dietary intervention,” *British Journal of Nutrition*, 2010, *104* (1), 125–134.
- Chang, Hung-Hao and David R Just**, “Health Information Availability and the Consumption of Eggs: Are Consumers Bayesians?,” *Journal of Agricultural and Resource Economics*, 2007, pp. 77–92.
- Chern, Wen S, Edna T Loehman, and Steven T Yen**, “Information, health risk beliefs, and the demand for fats and oils,” *The Review of Economics and Statistics*, 1995, pp. 555–564.
- Colditz, Graham A, JOANN E MANSON, and SUSAN E HANKINSON**, “The Nurses’ Health Study: 20-year contribution to the understanding of health among women,” *Journal of Women’s Health*, 1997, *6* (1), 49–62.
- Cutler, David M**, “Behavioral health interventions: what works and why,” *Critical perspectives on racial and ethnic differences in health in late life*, 2004, *643*, 674.
- , **Adriana Lleras-Muney, and Tom Vogl**, “Socioeconomic status and health: dimensions and mechanisms,” Technical Report, National Bureau of Economic Research 2008.
- **and** —, “Understanding differences in health behaviors by education,” *Journal of health economics*, 2010, *29* (1), 1–28.

- Einav, Liran, Ephraim Leibtag, and Aviv Nevo**, “Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?,” *Quantitative Marketing and Economics*, 2010.
- Finke, Michael S and Sandra J Huston**, “Factors affecting the probability of choosing a risky diet,” *Journal of Family and Economic Issues*, 2003, 24 (3), 291–303.
- Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams**, “Place-Based Drivers of Mortality: Evidence from Migration,” *Stanford University Working Paper*, 2018.
- Friel, Sharon, John Newell, and Cecily Kelleher**, “Who eats four or more servings of fruit and vegetables per day? Multivariate classification tree analysis of data from the 1998 Survey of Lifestyle, Attitudes and Nutrition in the Republic of Ireland,” *Public health nutrition*, 2005, 8 (2), 159–169.
- Goldman, Dana P and James P Smith**, “Can patient self-management help explain the SES health gradient?,” *Proceedings of the National Academy of Sciences*, 2002, 99 (16), 10929–10934.
- Greenland, Sander, James M Robins, and Judea Pearl**, “Confounding and collapsibility in causal inference,” *Statistical science*, 1999, pp. 29–46.
- Ippolito, Pauline M and Alan D Mathios**, “Information and advertising: The case of fat consumption in the United States,” *The American Economic Review*, 1995, 85 (2), 91–95.
- Kinnucan, Henry W, Hui Xiao, Chung-Jen Hsia, and John D Jackson**, “Effects of health information and generic advertising on US meat demand,” *American Journal of Agricultural Economics*, 1997, 79 (1), 13–23.
- Kirkpatrick, Sharon I, Kevin W Dodd, Jill Reedy, and Susan M Krebs-Smith**, “Income and race/ethnicity are associated with adherence to food-based dietary guidance among US adults and children,” *Journal of the Academy of Nutrition and Dietetics*, 2012, 112 (5), 624–635.
- Loftfield, Erikka, Marilyn C Cornelis, Neil Caporaso, Kai Yu, Rashmi Sinha, and Neal Freedman**, “Association of Coffee Drinking With Mortality by Genetic Variation in Caffeine Metabolism: Findings From the UK Biobank,” *JAMA internal medicine*, 2018, 178 (8), 1086–1097.
- Maki, Kevin C, Joanne L Slavin, Tia M Rains, and Penny M Kris-Etherton**, “Limitations of observational evidence: implications for evidence-based dietary recommendations,” *Advances in nutrition*, 2014, 5 (1), 7–15.
- Oster, Emily**, “Unobservable selection and coefficient stability: Theory and evidence,” *Journal of Business & Economic Statistics*, 2018, pp. 1–18.
- Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt**, “Poorly measured confounders are more useful on the left than on the right,” *Journal of Business & Economic Statistics*, 2018, (just-accepted), 1–34.
- Roosen, Jutta, Stéphan Marette, Sandrine Blanchemanche, and Philippe Verger**, “Does health information matter for modifying consumption? A field experiment measuring the impact of risk information on fish consumption,” *Review of Agricultural Economics*, 2009, 31 (1), 2–20.

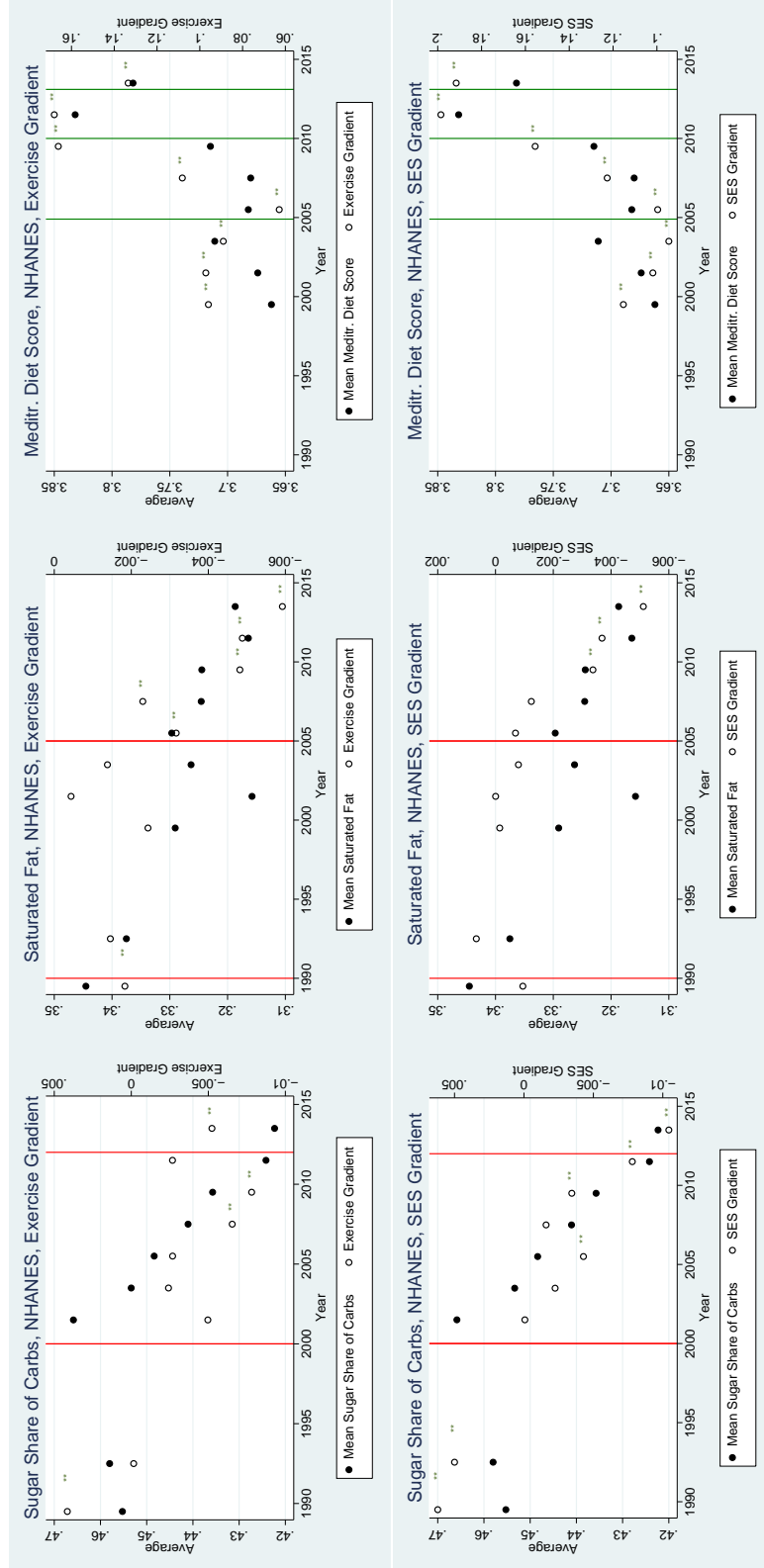
- Trichopoulou, Antonia, Tina Costacou, Christina Bamia, and Dimitrios Trichopoulos**, “Adherence to a Mediterranean diet and survival in a Greek population,” *New England Journal of Medicine*, 2003, *348* (26), 2599–2608.
- Vandenbroucke, Jan P, Erik Von Elm, Douglas G Altman, Peter C Gøtzsche, Cynthia D Mulrow, Stuart J Pocock, Charles Poole, James J Schlesselman, Matthias Egger, Strobe Initiative et al.**, “Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration,” *PLoS medicine*, 2007, *4* (10), e297.
- Yang, Quanhe, Zefeng Zhang, Edward W Gregg, W Dana Flanders, Robert Merritt, and Frank B Hu**, “Added sugar intake and cardiovascular diseases mortality among US adults,” *JAMA internal medicine*, 2014, *174* (4), 516–524.

Figure 1: Vitamin Consumption Levels and Other Behavior or Proxy Gradients



Notes: These figures show the co-movement between vitamin consumption or purchasing behavior and the exercise or socioeconomic gradient with respect to vitamin consumption or purchases over time. Green lines indicate events where the behavior was more recommended; red lines indicate changes to less recommended. Details of the events appear in Appendix Table A2. \*\*indicates significance at the 5% level.

Figure 2: Diet Behavior Levels and Other Behavior or Proxy Gradients



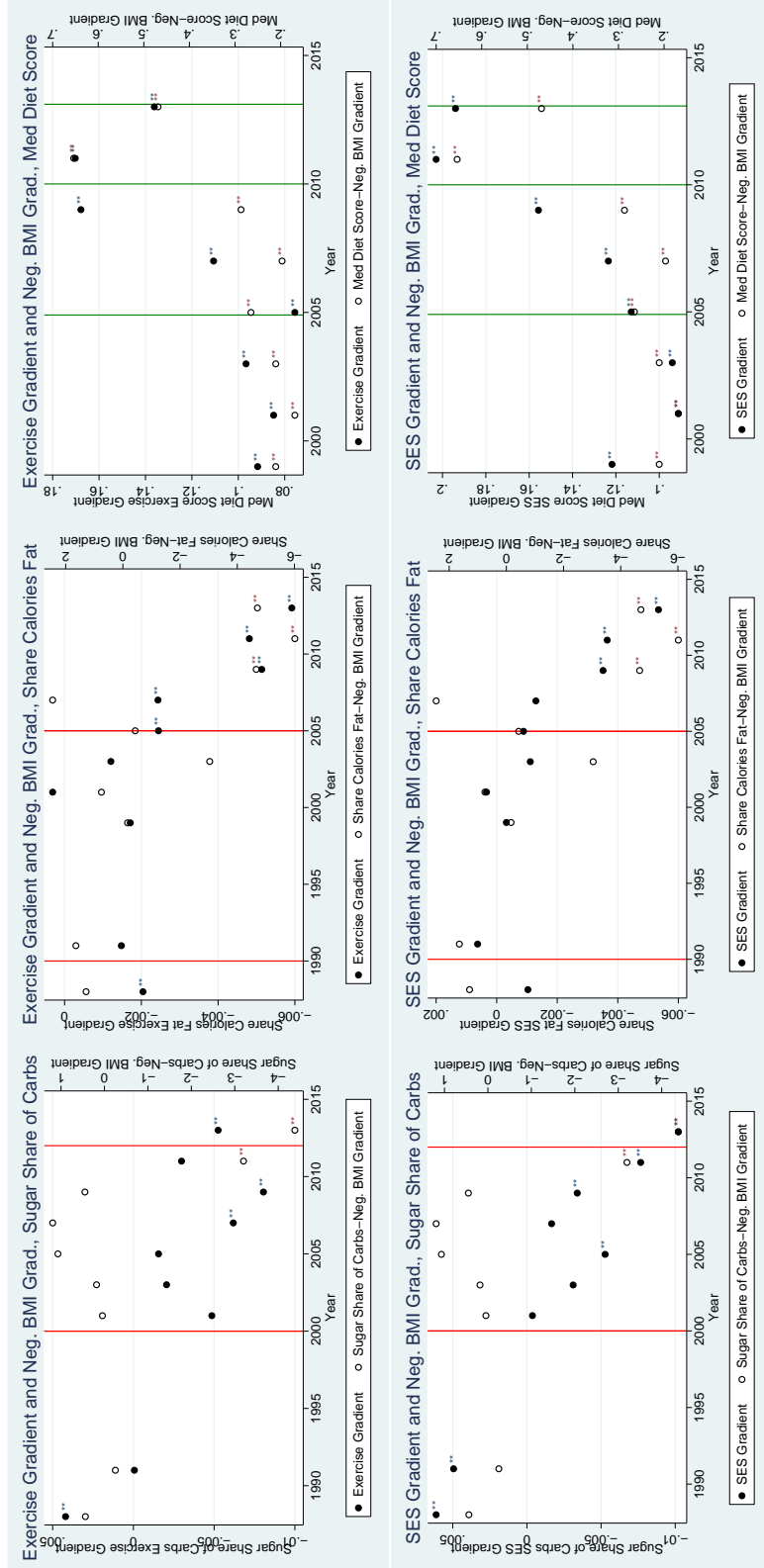
Notes: These figures show the co-movement between dietary behavior and the exercise or socioeconomic gradient with respect to dietary behavior over time. Green lines indicate events where the behavior was more recommended; red lines indicate changes to less recommended. Details of the events appear in Appendix Table A2. \*\*indicates significance at the 5% level.

Figure 3: Evidence from Publications on Vitamins



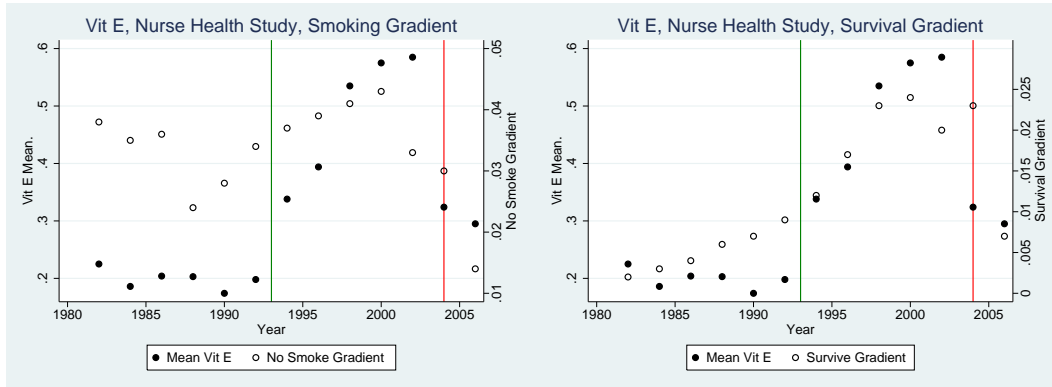
Notes: These figures show the share of significant vitamin-cancer relationships in published work using data from each period. Publications are identified from Pubmed searches and from published meta-analyses. The outcome (significant negative relationship between vitamin supplementation and cancer) is residualized with respect to the type of cancer. Studies with data which overlaps the time periods is assigned a partial weight in each time period.

Figure 4: Disease Gradients and Behavior/Proxy Gradients



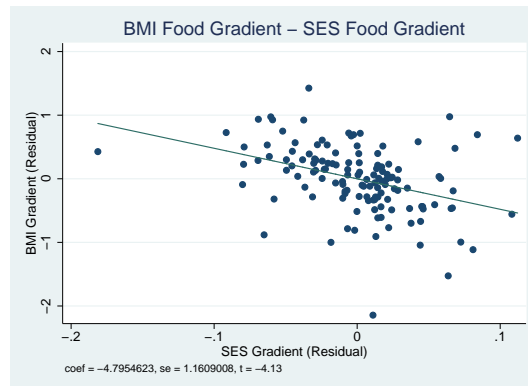
Notes: These figures show the co-movement between the Negative BMI gradient with respect to dietary behavior and the socioeconomic gradient with respect to dietary behavior over time. Green lines indicate events where the behavior was more recommended; red lines indicate changes to less recommended. Details of the events appear in Appendix Table A2. \*\*indicates significance at the 5% level.

Figure 5: Evidence from Nurse Health Study



Notes: These figures show co-movement between vitamin E supplementation rates the vitamin E-not smoking gradient (left figure) and the vitamin E-survival gradient (right figure). Data come from the Nurses Health Study. Regressions producing the gradients include controls for age fixed effects.

Figure 6: Disease Gradients and Behavior/Proxy Gradients: Individual Foods



Notes: This figure shows the relationship between the negative BMI gradient residualized with respect to individual foods fixed effects and the socioeconomic gradient residualized with respect to individual foods fixed effects. It shows the relationship between the BMI gradient and the socioeconomic gradient for each food and time period controlling for individual foods fixed effects. This figure is the graphical equivalent of the regression in column (1) of Table 4.



Table 1: **Vitamins Levels and Gradients: Co-Movements**

<b>Panel A: Vitamin Consumption, NHANES</b>				
	(1)	(2)	(3)	(4)
<i>Outcome:</i>	Exercise Gradient	SES Gradient	Smoking Gradient	Diet Quality Gradient
Take Vit Supplement	0.147*** (0.020)	0.269*** (0.023)	0.315*** (0.040)	0.573*** (0.192)
Supplement FE	YES	YES	YES	YES
Number of Obs.	20	20	20	18
<b>Panel B: Vitamin Consumption, HomeScan</b>				
	(1)	(2)		
<i>Outcome:</i>	SES Gradient	Smoking Gradient	Diet Quality Gradient	
Take Vit. Supplement	0.050 *** (0.013)	0.088 *** (0.024)	0.288 *** (0.070)	
Supplement FE	YES	YES	YES	
Number of Obs.	26	26	26	

Notes: This table shows statistical evidence on the co-movements between levels of behavior and the gradient with respect to various proxies for selection (socioeconomic status, exercise, medical care access and diet quality). Each cell represents a different regression of gradients on the level of behavior with Panel A focusing on vitamins in the NHANES and Panel B on vitamins in the HomeScan. All regressions include fixed effects for each behavior. There are fewer observations for diet quality as our measure cannot be constructed for 1999. \*indicates significance at the 10% level, \*\*indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.

Table 2: **Dietary Patterns Levels and Gradients: Co-Movements**

	(1)	(2)	(3)
<i>Outcome:</i>	Exercise Gradient	SES Gradient	Smoking Gradient
Diet Measure (Standardized)	0.099** (0.047)	0.192*** (0.049)	0.304** (0.114)
Behavior FE	YES	YES	YES
Number of Obs.	27	27	27

Notes: This table shows statistical evidence on the co-movements between levels of behavior and the gradient with respect to various proxies for selection (socioeconomic status, exercise, smoking). Each cell represents a different regression of gradients on the level of behavior. All regressions include fixed effects for each behavior. \*indicates significance at the 10% level, \*\*indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.

Table 3: **Co-movement: Behavior-Outcome Gradient and Selection Gradients**

<b>Panel A: Outcome Regression With No Controls</b>			
<i>Outcome:</i>	BMI Gradient	Obesity Gradient	Heart Health Gradient
Exercise Gradient [n=27]	-4.71*** (1.68)	-0.223** (0.106)	0.849*** (0.173)
SES Gradient [n=27]	-5.84*** (1.06)	-0.311*** (0.071)	0.617** (0.161)
Smoking Gradient [n=27]	-1.82** (0.675)	-0.096** (0.041)	0.130 (0.094)
<b>Panel B: Outcome Regression With Controls</b>			
<i>Outcome:</i>	BMI Gradient	Obesity Gradient	Heart Health Gradient
Exercise Gradient [n=27]	-3.68** (1.47)	-0.176* (0.096)	0.664*** (0.166)
SES Gradient [n=27]	-4.56*** (0.997)	-0.250*** (0.068)	0.434** (0.155)
Smoking Gradient [n=27]	-1.60*** (0.57)	-0.086** (0.036)	0.073 (0.084)
Diet Behavior FE	YES	YES	YES

Notes: This table shows statistical evidence on the co-movements between health gradients and selection gradients with respect to dietary behavior. The behavior-outcome regressions include controls for age and gender in all cases. In Panel B these also include controls for race, marital status, education and income. \* indicates significance at the 10% level, \*\* indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.

Table 4: **Co-movement Behavior-Outcome Gradient and Selection Gradients: Individual Foods**

	(1)	(2)	(3)
<i>Outcome:</i>	BMI Gradient	Obesity Gradient	Heart Health Gradient
Exercise Gradient [n=132]	-4.358*** (1.44)	-0.194 (0.129)	-0.174 (0.207)
SES Gradient [n=132]	-4.79*** (1.16)	-0.167 (0.156)	0.003 (0.249)
Smoking [n=132]	-2.28*** (0.53)	-0.151*** (0.059)	0.313*** (0.091)
Foods FE	YES	YES	YES

Notes: This table shows statistical evidence on the co-movements between health gradients and selection gradients with respect to food consumption. The behavior-outcome regressions include controls for age and gender. The exercise, SES and doctor gradient regression are run using a Probit model. \* indicates significance at the 10% level, \*\* indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.

Table 5: **Smoking, Death and Vitamin E: Nurse Health Study**

<b>Panel A: Smoking Behavior by Vitamin E Status, 1990-1992 vs. 1998-2000</b>			
		Vit. E Status, 1998-2000	
		<i>Yes</i>	<i>No</i>
Vit. E Status, 1990-92	<i>Yes</i>	0.129 [N=14,595]	0.141 [N=834]
	<i>No</i>	0.137 [N=19,479]	0.187 [N=12,914]

<b>Panel B: Death Rates by Vitamin E Status, 1990-1992 vs. 1998-2000</b>			
		Vit. E Status, 1998-2000	
		<i>Yes</i>	<i>No</i>
Vit. E Status, 1990-92	<i>Yes</i>	0.122 [N=14,765]	0.115 [N=846]
	<i>No</i>	0.101 [N=19,464]	0.104 [N=12,937]

Notes: This table shows smoking behavior by Vitamin E status, around the adoption and dis-adoption of Vitamin E, in the Nurse Health Study. Panel A looks at smoking behavior in 1990/1992 by group, based on their use of Vitamin E in the earlier and later period. Panel B look at death rates in the ten years following the second period for these same groups. Good news about Vitamin E was released in 1993, and bad news in 2004.

Table 6: **Proportional Selection Adjustment Results**

<i>Outcome/Treatment</i>	Period 1 Coeff.	Period 2 Coeff.	Period 3 Coeff.	<b>Implied <math>\beta</math></b>	Implied $\delta$
BMI/Sugar	-0.043	-0.66**	3.18***	<b>0.143</b>	[0.07,0.16]
Obesity/Sugar	-0.006	-0.052**	0.17***	<b>-0.003</b>	[0.07,0.09]
Heart Health/Sugar	-0.290***	-0.528***	-0.549***	<b>-0.495***</b>	[0.014,0.05]
BMI/Fat	-1.12***	4.47***		<b>2.07</b>	0.043
Obesity/Fat	-0.022	0.242***		<b>0.148</b>	0.026
Heart Health/Fat	-0.369***	-0.967***		<b>2.5***</b>	0.33
BMI/Med. Diet	-0.230***	-0.496***		<b>-0.259***</b>	0.031
Obesity/Med Diet	-0.015***	-0.032***		<b>-0.131***</b>	0.27
Heart Health/Med Diet	0.030***	0.036***		<b>0.017***</b>	0.024

Notes: This table shows estimates of causal impacts in these relationships under an assumption of proportional selection between observed and unobserved variables (Oster, 2018). In each case we solve for the implied  $\delta$  and causal impact  $\beta$  under the assumption of proportional selection. The calculations are made using a grid search; details are in Section 6. Standard errors on  $\beta$  are based on a bootstrap with 1000 replications. \*indicates significance at the 10% level, \*\*indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.

## Appendix A: Figures and Tables

Table A1: Data Source Summary

<b>Panel A: Selection into Behavior</b>				
Domain	Dataset	Behavior	Other Health Behaviors	Demographic Health Confounds
Vitamins	NHANES	Intake of Vitamin D and Vitamin E	Exercise, Access to medical care	Income, Education
Vitamins	Nielsen HomeScan	Purchases of Vitamin D and Vitamin E	Smoking	Income, Education
Vitamin E	Nurses Health Study	Intake of Vitamin E	Smoking	
Diet Patterns	NHANES	Sugar consumption, Fat Consumption, Adherence to Mediterranean Diet	Exercise, Smoking	Income, Education
<b>Panel B: Behavior Effect on Disease</b>				
Domain	Dataset	Behavior	Outcomes	
Vitamins	Publications	Intake of Vitamin D and Vitamin E	Cancer	
Vitamins	Nurses Health Study	Intake of Vitamin E	Death	
Diet Patterns	NHANES	Sugar consumption, Fat Consumption, Adherence to Mediterranean Diet	BMI, Obesity, Heart Health	
Individual Foods	NHANES	Consumption of particular foods	BMI, Obesity, Heart Health	

Notes: This table summarizes the data used in the paper.

Table A2: Information Events

<i>Behavior</i>	<i>Event 1</i>	<i>Event 2</i>	<i>Event 3</i>
<i>Vitamin C Supplementation</i>	<b>1998, Negative:</b> <i>Nature</i> article suggesting vitamin C has pro-oxidant properties; is dangerous. Corresponding NYTimes coverage		
<i>Vitamin D Supplementation</i>	<b>2007, Positive:</b> Several studies, NEJM summary piece, NY Times coverage suggest Vitamin D good for health (cancer, fractures, etc). Corresponding growth in Google Trends.	<b>2011/2012, Negative:</b> IOM report suggests Vitamin D overblown, corresponding summary articles, coverage in NY Times. Additional studies in 2012 with similar findings. Google Trends stagnation.	
<i>Vitamin E Supplementation</i>	<b>1993, Positive:</b> Two studies in NEJM report reduction in heart disease for both men and women with use of Vitamin E supplements	<b>2004, Negative:</b> Widely covered meta-analysis of Vitamin E shows high doses increase mortality. Large Google trends spike.	
<i>Sugar in Diet</i>	<b>2000, Negative:</b> First explicit mention in US Dietary guidelines of avoidance of added sugars.	<b>2011/2012, Negative:</b> Extensive media coverage of health costs of sugar; “toxic sugar” in NY Times and 60 Minutes Segment.	<b>2015, Negative:</b> First explicit restriction on sugar share in US Dietary Guidelines.
<i>Saturated Fat</i>	<b>1990, Negative:</b> First explicit restriction on saturated fat share in US dietary guidelines (<10%)	<b>2005, Negative:</b> Further restrict saturated fat to <7% for people with heart disease.	<b>2015, Positive:</b> US Dietary guidelines return to 10% restriction for everyone.
<i>Mediterranean Diet</i>	<b>2004, Positive:</b> Two JAMA articles show health benefits of Mediterranean diet. Google trends spike.	<b>2009/2010, Positive:</b> Series of articles on role of Mediterranean diet in addressing cognitive decline. Google trends spike.	<b>2013, Positive:</b> Large randomized trial shows mortality reductions from Mediterranean Diet. Heavily covered in media. Very large Google trend spike.

Notes: This table shows the information events identified for each outcome. Events were identified by searching for well-cited publications, media coverage and Google search spikes.

Table A3: Range of Estimates for Health Impacts of Diet

<b>Panel A: Impacts of Sugar</b>				
<i>Outcome</i>	<i>Overall Effect</i>	<i>Minimum Effect</i>	<i>Maximum Effect</i>	<i>p-value, equal effects</i>
BMI	-0.030 (0.208)	-1.19 (0.79) [2007]	4.38*** (0.84)[2013]	0.000
Obese (0/1)	-0.012 (0.015)	-0.140** (0.058)[2007]	0.221*** (0.057)[2013]	0.000
Heart Health	-0.45*** (0.036)	-0.690*** (0.123)[2009]	-0.113 (0.080)[1988]	0.000
<b>Panel B: Impact of Fat</b>				
<i>Outcome</i>	<i>Overall Effect</i>	<i>Minimum Effect</i>	<i>Maximum Effect</i>	<i>p-value, equal effects</i>
BMI	-0.71 (0.41 )	-2.45 (1.64)[2007]	6.01*** (1.66)[2011]	0.000
Obese (0/1)	-0.019 (0.030 )	-0.159 (0.122)[2007]	0.317*** (0.129)[2003]	0.000
Heart Health	-0.781*** (0.072)	-1.22*** (0.30)[2005]	0.202 (0.166)[1991]	0.000
<b>Panel C: Impact of Mediterranean Diet</b>				
<i>Outcome</i>	<i>Overall Effect</i>	<i>Minimum Effect</i>	<i>Maximum Effect</i>	<i>p-value, equal effects</i>
BMI	-0.31*** (0.024)	-0.653*** (0.067)[2011]	-0.168*** (0.064)[2001]	0.000
Obese (0/1)	-0.020*** (0.002)	-0.038*** (0.005)[2011]	-0.009** (0.005)[2001]	0.000
Heart Health	0.045*** (0.0041)	0.014 (0.012)[2003]	0.071*** (0.011)[2009]	0.000

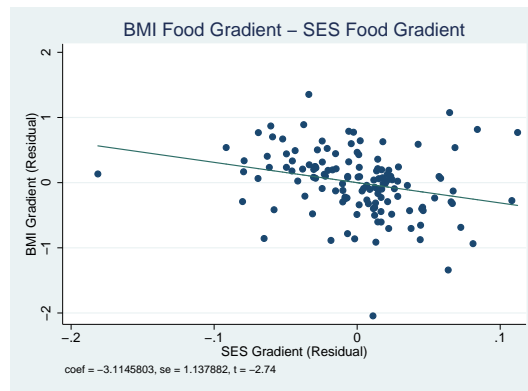
Notes: This table shows the range of estimates of the impact of behavior on health across years in the NHANES. Panel A focuses on sugar consumption, Panel B focuses on fat consumption and Panel C focuses on adherence to a Mediterranean-style diet. The first column reports the overall effect of behavior on health from a regression with data pooled across all years. The second and third column report, respectively, the minimum and maximum effect of behavior on health over time, with the corresponding year reported in square brackets. Each cell represents a different regression and all regressions include controls for age and gender. The fourth column reports the p-value for a test of equality of coefficients across all years. \* indicates significance at the 10% level, \*\* indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.

Table A4: **Individual Food Results, Disease Regressions Include Controls**

	(1)	(2)	(3)
<i>Outcome:</i>	BMI Gradient	Obesity Gradient	Heart Health Gradient
Exercise Gradient [n=132]	-3.11*** (1.13)	-0.102 (0.152)	-0.233 (0.240)
SES Gradient [n=132]	-3.41** (1.37)	-0.089 (0.127)	-0.431** (0.196)
Smoking [n=132]	-1.93*** (0.51)	-0.136** (0.057)	0.250*** (0.090)
Foods FE	YES	YES	YES

Notes: This table shows statistical evidence on the co-movements between health gradients and selection gradients with respect to food consumption. This echoes Table 4 in the paper, but with additional demographic controls included in the behavior-outcome regressions. \* indicates significance at the 10% level, \*\* indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.

Figure A1: **Individual Foods Analysis, Disease Regressions Include Controls**



Notes: This figure show the relationship between the BMI gradient residualized with respect to individual foods fixed effects and the socioeconomic gradient residualized with respect to individual foods fixed effects. It shows the relationship between the BMI gradient and the socioeconomic gradient for each food and time period controlling for individual foods fixed effects. This figure echoes Figure 6 in the main text, but the behavior-outcome regressions are estimated with additional demographic controls.

## Appendix B: Proofs

### Proof of Proposition 1.

This follows trivially from observing that the rank of  $A$  in the set of behaviors improves as a result of the change in recommendation and it is therefore weakly more likely to be adopted by each individual.

### Proof of Proposition 2.

First of all, note that, after the change in recommendation  $h^A$  decreases as a consequence of behavior  $A$  improving its position in the overall ranking of behaviors. This implies that, by Proposition 1, more people will start adopting behavior  $A$ . Because of the mapping between  $h_i^*$  and the individual's health values, those who will adopt it are those with relatively higher  $h_i^*$ .

Now let  $Cov_t(A, h^*)$  denote the covariance between  $A$  and  $h^*$  before the change in recommendation and  $Cov_{t+1}(A, h^*)$  denote the same covariance after the change in recommendation. We want to show that:

$$Cov_{t+1}(A, h^*) \geq Cov_t(A, h^*)$$

First of all, note that we can write Assumption 1 in the following way:

$$\frac{E_{t+1} [h^* | h^* > h_{t+1}^A] - E_{t+1} [h^*]}{E_t [h^* | h^* > h_t^A] - E_t [h^*]} \geq \frac{P(h^* > h_t^A)}{P(h^* > h_{t+1}^A)}$$

Note that  $E[h^*]$  is constant across periods, so we can write:

$$\frac{E_{t+1} [h^* | h^* > h_{t+1}^A] - E[h^*]}{E_t [h^* | h^* > h_t^A] - E[h^*]} \geq \frac{P(h^* > h_t^A)}{P(h^* > h_{t+1}^A)}$$

Rearranging we have that:

$$\begin{aligned} P(h^* > h_{t+1}^A) (E_{t+1} [h^* | h^* > h_{t+1}^A] - E[h^*]) &\geq P(h^* > h_t^A) (E_t [h^* | h^* > h_t^A] - E[h^*]) \\ P(h^* > h_{t+1}^A) E_{t+1} [h^* | h^* > h_{t+1}^A] - P(h^* > h_{t+1}^A) E[h^*] &\geq P(h^* > h_t^A) E_t [h^* | h^* > h_t^A] - P(h^* > h_t^A) E[h^*] \end{aligned}$$

Using the definition of conditional expectation and the property of indicator variables for which  $P(h^* > h^A) = E[1(h^* > h^A)]$ , we can then write:

$$E_{t+1} [h^* 1(h^* > h_{t+1}^A)] - E_{t+1} [1(h^* > h_{t+1}^A)] \geq E_t [h^* 1(h^* > h_t^A)] - E_t [1(h^* > h_t^A)]$$

Note that defining  $A = 1(h_i^* > h_t^A)$ , by the usual characterization of the covariance, the previous expression gives exactly the result we were looking for:

$$Cov_{t+1}(A, h^*) \geq Cov_t(A, h^*)$$

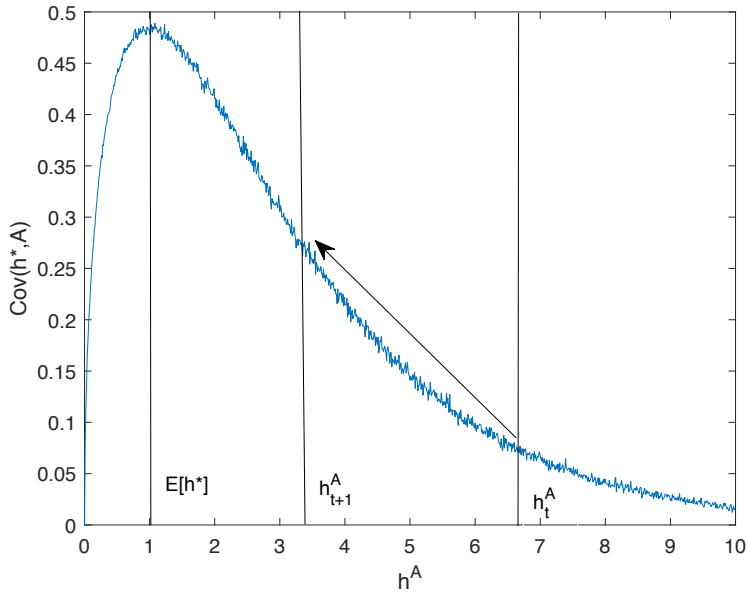
We now develop some intuition for this, under particular distributional assumptions.

The two figures below show two simulation examples of change in recommendation where Assumption 1 holds and as a consequence Proposition 2 is true. Figure B1 depicts an example where  $h^*$  is distributed as a chi-squared with degree 1 and behavior  $A$  is relatively unusual. In particular,  $h_t^A$  is in the right tail of the distribution of  $h^*$  and this implies that the adopters before the change in recommendation, that is those such that  $h_i^* \geq h_t^A$ , are relatively few. When behavior  $A$  is relatively unusual, even if the change in  $\kappa_A$  is big as depicted in Figure 1, Proposition 2 holds and the covariance between  $h^*$  and  $A$  increases after the change in recommendation.

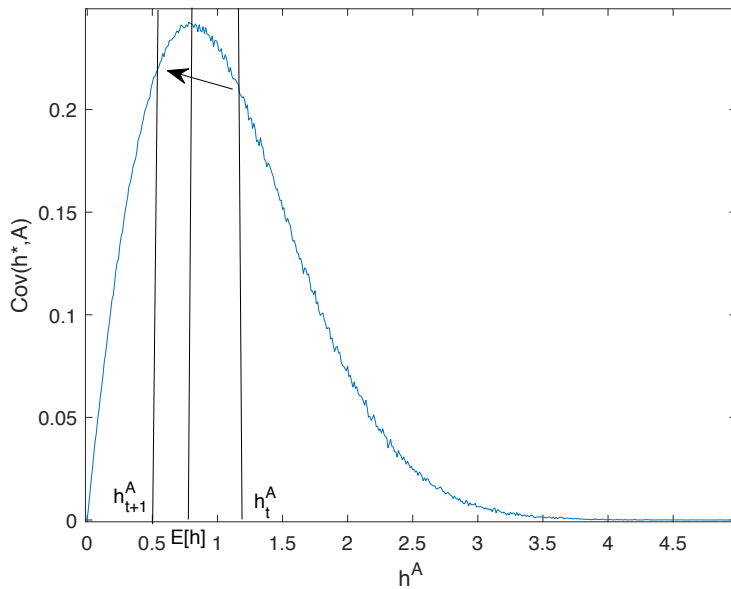
Figure B2 shows a different example where  $h^*$  is distributed as a standard normal truncated at zero. In this case, behavior  $A$  is relatively common before the change in recommendation ( $h_t^A$  is just above the average  $E[h]$ ) but the change in recommendation is such that  $\kappa_A$  does not increase much. As it can be seen from the Figure, if the change in recommendation is not big then Proposition 2 will hold also when behavior  $A$  is relatively common.



**Figure B1:** Example of change in recommendation when  $h^*$  is distributed as a chi-squared of degree 1 and behavior  $A$  is unusual



**Figure B2:** Example of change in recommendation when  $h^*$  is distributed as a standard normal truncated at zero and behavior  $A$  is relatively common



### Proof of Proposition 3

Under Assumption 1, adoption of  $A$  is weakly increasing in  $h^*$  and  $h^*$  is increasing in  $\omega$ . By composition of increasing functions, it follows that the adoption of  $A$  is increasing in  $\omega$ .

**Proof of Proposition 4:**

The true model is  $Y_i = \alpha + \vartheta\Lambda_i + \epsilon_i$ . We rewrite this to pull out  $A$  specifically and denote the modified  $\Lambda$  vector and coefficient vector with primes:

$$Y_i = \alpha + \beta A_i + \vartheta \mathring{\Lambda}_i + \epsilon_i$$

Consider first the case where we observe no elements of  $\Lambda$ . The feasible regression is then

$$Y_i = \alpha + \mathring{\beta} A_i + \mathring{\epsilon}_i.$$

By a straightforward application of the omitted variable bias formula, we can write

$$\mathring{\beta} = \beta + \vartheta \frac{Cov(A, \mathring{\Lambda})}{Var(A)}$$

Trivially, this bias is increasing as  $Cov(A, \Lambda)$  increases.

Now consider a case where we observe a subset of behaviors such that we can write  $\mathring{\Lambda} = \Delta_1 \widehat{\Lambda} + \Delta_2 \widetilde{\Lambda}$  where  $\widehat{\Lambda}$  is observed and  $\widetilde{\Lambda}$  is not, and define this such that the two components are orthogonal. We can now estimate

$$Y_i = \alpha + \widehat{\beta} A_i + \widehat{\Delta}_1 \widehat{\Lambda}_i + \mathring{\epsilon}_i.$$

Again, by the omitted variable bias formula, we have

$$\widehat{\beta} = \beta + \Delta_2 \frac{Cov(A, \widetilde{\Lambda})}{Var(A)}$$

Which is also increasing in the covariance between  $A$  and  $\Lambda$  since  $Cov(A, \widetilde{\Lambda})$  is an increasing function of  $Cov(A, \Lambda)$  by the linearity of covariance.

**Proof of Corollary 1:**

This follows trivially from the results above.

## Appendix C: List of Foods for Individual Food Analysis

Individual Foods selected for the Food-Specific Analysis

<i>Foods Selected</i>	Period 1	Period 2	Period 3	Period 4	Trend
Banana	-0.245	-0.735***	-0.843***	-0.933***	-0.473***
Berries	-0.642**	-1.498***	-1.266***	-1.630***	-0.322***
Biscuits	0.256	0.640**	0.816***	1.179***	0.354***
Brown Rice	0.040	-0.991	-1.204*	-1.985***	-0.785***
Cheese	-0.397**	-0.382**	-0.521***	0.163	0.075
Cooked Cereals	-0.649	-1.307	-0.864	-4.661***	-1.179***
Corn	-0.557***	-0.188	0.326	0.224	0.354***
Crackers	-0.215	-0.172	-0.789***	-0.145	-0.072
Egg Soup	0.564	-0.161	0.397***	-3.598	-1.540**
Eggs (white)	0.672	-1.324*	-1.265**	-1.176**	-0.596**
Egg (whole)	0.357**	0.490**	0.063	-0.100	-0.212***
Fruit Juice	0.0783	0.145	0.055	0.636***	0.140**
Ice Cream	-0.564***	-0.243	-0.205	0.721***	0.337***
Instant Soup	0.989*	0.460	0.544	-1.076**	-0.632***
Legume Soup	-1.318**	-1.597***	-1.220**	-3.160***	-0.608**
Lunch Sandwich	0.098	0.672***	0.764***	1.725***	0.469***
Margarine	-0.386***	-0.170	0.096	0.585***	0.387***
Nuts	-0.748***	-0.558***	-1.397***	-1.606***	-0.410***
Other Milk	-0.266	-0.474	-0.488	-1.915***	-0.491***
Peanut Butter	-0.642***	-0.845***	-0.967***	-1.498***	-0.283***
Peas and Lentils Mix	-0.337	-1.191**	-1.339**	-1.719***	-0.467**
Processed Cheese	0.104	0.427**	0.558***	0.937***	0.278***
Rye Bread	-0.445	-1.034***	0.472	0.951	0.557***
Salad Dressing	-0.357***	-0.007	-0.174	0.208	0.144**
Seafood	-0.192	-0.105	-0.440***	-0.880***	-0.226***
Seeds	0.339	0.080	-1.487***	-1.393***	-0.743***
Soda	0.736***	0.859***	0.883***	1.399***	0.217***
Soybean Products	-0.844**	-1.710***	-1.498***	-2.190***	-0.404***
Tortilla	0.546***	0.113	0.061	1.039***	0.081
Vegetable Mixture	0.060	0.057	0.022	-1.144***	-0.374***
Vegetable Soup	-0.348	-0.700**	-0.813***	-1.734***	-0.353***
White Rice	-0.108	-0.606	0.067	-2.126***	-0.710***
Yellow Vegetables	-0.196	-0.726**	-0.681**	-1.529***	-0.379***

*Notes:* this table lists all the individual foods selected for the food specific analysis along with the coefficients of their effects on BMI in each time period and the coefficient of the linear trend effect. Cooked cereals include cream of wheat and cooked barley but not oatmeal. Other milk includes buttermilk, sweetened condensed milk and evaporated milk. Yellow vegetables include squashes, gourds and pumpkins. Peas and lentils mix refers to any vegetable mixture that has peas and lentils in it \* indicates significance at the 10% level, \*\* indicates significance at the 5% level, \*\*\* indicates significance at the 1% level.