J Popul Econ (2019) 32:877–913
https://doi.org/10.1007/s00148-018-0709-9

ORIGINAL PAPER

Body weight and Internet access: evidence from the rollout of broadband providers

Michael DiNardi[1] · **Melanie Guldi**[2] · **David Simon**[3]

Received: 29 January 2017 / Accepted: 7 June 2018 / Published online: 27 June 2018

Abstract Obesity has become an increasingly important public health issue in the USA and many other countries. Hypothesized causes for this increase include declining relative cost of food and a decreasing share of the population working in labor-intensive occupations. In this paper, we suggest another factor: the Internet. Increasing Internet access could affect body weight through several channels. First, more time spent using the Internet, a sedentary activity, could lead to increases in body weight. Second, the prior literature has shown that economic activity (and income) increase with Internet access: given a positive health-income gradient, obesity rates could likewise increase, although the empirical evidence on the income-obesity gradient is mixed. Third, the Internet increases information and creates the possibility for online peer networks. Theoretically, increases in information should lead to more optimal consumer choices. At the same time, greater networking opportunities may result in peers having greater influence over positive or negative health behaviors. While we are unable to fully test these mechanisms, we are able to use the rollout of broadband Internet providers as a plausibly exogenous source of variation in Internet access to identify the reduced form effect of Internet use on body weight. We show that greater broadband coverage

Responsible editor: Erdal Tekin

✉ David Simon
david.simon@uconn.edu

Michael DiNardi
michael.dinardi@uconn.edu

Melanie Guldi
mguldi@ucf.edu

[1] Department of Economics, University of Connecticut, 365 Fairfield Way, U-1063, Storrs, CT 06269-1063, USA

[2] Department of Economics, University of Central Florida, 4336 Scorpius Street, Orlando, FL 32816-1400, USA

[3] Department of Economics, University of Connecticut and NBER, 365 Fairfield Way, U-1063, Storrs, CT 06269-1063, USA

increases the body weight of white women and has both positive and negative effects on modifiable adult health behaviors including exercise, smoking, and drinking.

Keywords Obesity · Exercise · Health · Information

JEL Classification I12 · I18 · D8

1 Introduction

Economists have long hypothesized that information is an important part of choice theory (Stigler 1961) and it is reasonable to assume that increases in access to information brought about by the Internet, which has transformed the way consumers acquire information, would improve health decisions and consequently health outcomes. However, this is not necessarily the case: consumers may substitute information from the Internet for visits to health professionals, resulting in exposure to lower quality health information (Wagner et al. 2001).[1] Beyond a direct information channel, the increasing prevalence of the Internet could influence health behaviors in other ways. Since it is a static activity, more time allocated toward surfing the Internet could increase body weight and this may lead individuals to experience the negative health consequences of a sedentary lifestyle (Owen et al. 2010). Additionally, social connections made through the Internet could influence weight gain and health through peer effects. Alternatively, increased Internet availability in an area may increase income which in turn could affect health given the positive health-income gradient, although the empirical evidence on the income-obesity gradient is mixed. In this paper, we explore the reduced form relationship between Internet access and adult body weight.

We use the rollout of broadband providers across counties in the USA during the 2000s as a plausibly exogenous proxy for increasing Internet access and use over this period. Similar identification strategies relying on the rollout of a policy or new technology have been used to estimate the effects of the food stamp program (Almond et al. 2011; Hoynes and Schanzenbach 2009), the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) (Hoynes et al. 2011), community health centers (Bailey and Goodman-Bacon 2015), and, of particular relevance to this work, electricity (Bailey and Collins 2011), broadcast television (Gentzkow 2006), and broadband Internet (Bhuller et al. 2013; Bellou 2015; Guldi and Herbst 2017; Kolko 2012). During the period of broadband rollout, the consumption of information on Web MD increased 20 fold, suggesting a relationship between broadband availability and the consumption of online health information. An existing body of work examines the relationship between increased use of the Internet to search for health information and the demand for health services,[2] but less research focuses on whether Internet availability directly affects health outcomes. The research that has been conducted is limited by time period, sample size, or is purely descriptive.[3] Given

[1] Additionally, without expert guidance, the large quantity of information available could lead consumers to accidently misuse the information they do receive.

[2] We describe this work in more detail in the Background section.

[3] For example, Bessière et al. (2010) use a random sample of the US population, but their study only covers 2 years.

that the Internet is increasingly used as a clearinghouse for health information[4] and that this appears to have accelerated with the advent of widespread high-speed access, understanding whether broadband access affects health outcomes is of great importance to understanding aggregate public health trends.

To investigate the relationship between Internet and health, we focus on body weight, overweight status, and obesity as our primary health outcomes. Obesity and the rising share of overweight individuals is an increasing public health concern in many countries including the USA (Cawley and Meyerhoefer 2012). However, the social causes behind this increase in body weight are not entirely understood. Additionally, weight is a health outcome modifiable through behavior: time use, diet, exercise, and the utilization of weight loss products all directly influence BMI. Likewise, consumers' decisions related to these products are all potentially affected by Internet availability. While much research has been devoted to understanding weight gain over time, little work has examined the relationship between technology and body weight.[5] We also look at modifiable health behaviors including exercise, smoking, and drinking to provide evidence on the pathways by which Internet access affects weight gain. Ours is the first paper we are aware of to directly estimate the reduced form effect of the rollout of broadband service on health. We find that the expansion of broadband coverage is associated with increases in the average BMI and obesity of women and that these effects are particularly salient for white women. These results suggest technology is an important part of the obesity discussion.

2 Conceptual framework

We suggest several channels through which the Internet may influence an individual's health behavior. Individuals are assumed to be making rational choices to optimize their utility subject to income and time constraints but with imperfect information. We posit that increases in access to the Internet may lead to increases in a consumer's information set, expansion of the consumer's peer network, changes in time use, and increases the consumer's income. These may in turn influence the individual's allocation of time for sedentary or dynamic activities, which may in turn affect measured body weight.

2.1 Information channel

Access to online health information has grown over time. Use of the Internet as a key source of this information has become increasingly common as consumers turn to websites, discussion boards, and social media. For example, the number of unique visitors to WebMD, an online publisher of health information, increased from 1.7

[4] Amante et al. (2015) provide evidence that individuals search for health information online, especially when it is difficult to access this information from health care providers.

[5] There is a growing body of work examining the effectiveness of smart phone applications and wearable technology (for example, Jakicic et al. 2016). These technologies, however, were largely developed after the period we consider.

million in December 1999 to a monthly average of 40.8 million in the third quarter of 2007.[6] Data from the Pew Research Center's Internet and American Life Project shows Internet use among adults increased from 46% in March 2000 to 75% in December of 2007. Over the same period, the share of Internet users who ever looked for health information online increased from 55 to 75%. Broken down by gender, 61% of women and 47% of men reported using the Internet to look for health information in 2000 and this increased to 81% of women and 68% of men by 2007, suggesting that women may engage with online health information more readily than men.[7] Quality of the health information discovered online, however, is variable. At one end of the quality spectrum are sources that provide correct and timely information. At the other end, information may be inaccurate, misleading, and potentially dangerous (Impicciatore et al. 1997; Akatsu and Kuffner 1998; Donald et al. 1998; McLellan 1998; Biermann et al. 1999; Purcell, Wilson, and Delamothe, Purcell et al. 2002).[8,9] Poor health information, or overwhelming amounts of information,[10] may lead consumers to forgo visits to professionals when more reliable advice is in fact needed. With this in mind, there are opposing viewpoints regarding how online information influences the consumption of health services. Some researchers find online health information is a complement to health services (Suiziedelyte 2012), and others find it serves as a substitute (Wagner et al. 2001). Although individuals may use online resources with the intent to make informed health-focused lifestyle changes, the potential difficulty in assessing the quality of this information may curb their ability to improve own health.

2.2 Social networks and time use

Beyond direct health information effects, the use of the Internet may also affect health through social networks. Networking through the Internet could alter health-related behaviors that have been shown to be influenced by peer effects including: positive health behaviors such as exercising (Carrell et al. 2011); or negative behaviors such as drinking alcohol, smoking, or using illegal drugs (Kremer and Levy 2008, b; Lundborg 2006).[11]

[6] http://investor.shareholder.com/wbmd/releasedetail.cfm?releaseid=249537&CompanyID=HLTH and http://investor.shareholder.com/wbmd/releasedetail.cfm?releaseid=274852&CompanyID=WBMD, accessed April 26, 2018.

[7] Data available from http://www.pewInternet.org/files/2014/01/Usage-Over-Time-_May-2013.xlsx. Accessed June 20, 2016.

[8] The issue of health information quality is so pervasive that the US National Institutes of Health has a webpage with resources to help consumers evaluate the quality of health-related websites. See https://nccih.nih.gov/health/webresources. Accessed August 18, 2016.

[9] The issue of quality is particularly salient in the work of Culver et al. (1997), who analyze messages from an online medical discussion group. They find 89% of the messages were authored by users without professional training, one third of the messages were inconsistent with conventional medical practices, and only 9% of the medical information provided by those without professional training contained a published citation. Similarly, Biermann et al. (1999) find 35% of websites with medical information about Ewing's sarcoma did not contain peer-reviewed sources, and some pages contained incorrect or misleading information.

[10] Some argue there is a glut of disorganized health-related information online (Donald et al. 1998; Berland et al. 2001; Purcell et al. 2002).

[11] Additionally, the Internet facilitates illegal drug transactions via the "dark web." http://www.newsweek.com/drugs-dark-web-silk-road-488957. Accessed October 13, 2016.

Additionally, when deciding what activities to engage in, individuals face a time constraint. Using the Internet is a fairly sedentary activity. Whether increasing time spent on the Internet leads to weight gain is in part a question of what type of activity it displaces. If more time is spent online at the expense of time spent reading a book, then we might not expect much change in body weight since both activities are similarly sedentary. If, however, an individual previously walked around a library looking for information and subsequently searches for this information online, we expect increases in body weight, all other things equal.

2.3 Income channel

An earlier literature has shown that Internet access improves wages, productivity, and growth (Kolko 2012; Akerman et al. 2015; Atasoy 2013). An expansive literature in economics has documented a positive relationship between income and health (e.g., Case and Paxon 2008). Increases in income is another mechanism by which Internet access could affect body weight. While the income-health gradient is generally positive, a number of studies show that the relationship between obesity and economic conditions is quite complicated, suggesting that the income effects of the Internet on health could be positive or negative.[12]

3 Background

3.1 Media, Internet access, and socioeconomic outcomes

There is a rich literature documenting the influence of media on socioeconomic outcomes. Researchers have found evidence that the introduction of broadcast television to a market leads to drops in voter turnout (Gentzkow 2006), and improvements in test scores (Gentzkow and Shapiro 2008). Others have found that the variety of television programming can influence political outcomes (DellaVigna and Kaplan 2007), fertility (Jensen and Oster 2009; La Ferrara, Chong, and Duryea 2012; Kearney and Levine 2015a; and Trudeau 2015), and can affect child outcomes (Kearney and Levine 2015b). Other researchers have examined the increasing availability of high-speed Internet and found it to be associated with improvements in wages and labor market opportunities (Akerman et al. 2015; Atasoy 2013; Dettling 2017; Kolko 2012). Additionally, the rollout of broadband has been linked with a wide variety of other outcomes including increases in voter turn-out (Poy and Schuller 2016); marriage market matching (Bellou 2015; Potarca 2017); reductions in teen fertility (Guldi and Herbst 2017); and increased incidence and reporting of sex crimes (Bhuller et al. 2013). Complementing this other work, some researchers have explored the relationship between the Internet and health outcomes (Bessière et al. 2010) or the demand for health care services (Baker et al. 2003; Suziedelyte 2012; Wagner et al.

[12] The cross-sectional relationship suggests that higher income is associated with lower levels of obesity. However, economic recessions have been known to reduce body weight in the severely obese (Ruhm 2005). Similarly, income transfers to low-income Native American adults through a casino opening increased obesity in their children (Akee et al. 2015).

2001). Yet, no paper has examined the potentially causal relationship between broadband expansion and obesity, a prominent health concern in many countries.

3.2 Obesity

Since 1980, the world obesity rate has doubled and today most of the world's population live in countries where being obese is more likely to cause death than being underweight.[13] Figure 1 demonstrates the trends in the rates of obesity and overweight for US adults aged 18 to 64, separately by race and gender, from 1990 to 2007. There are several patterns evident in the raw data. First, rates of obesity and overweight are lower for white women than non-white women while they are more similar for men. Second, the obesity and overweight rates rose sharply across all race and gender subgroups, affecting a large fraction of the population. By 2007, half or more of each group is classified as overweight. In response to these alarming trends, obesity has been declared as one of the leading problems in public health in the USA and other developed countries. The health costs of obesity are estimated to be at least 9.1%, or as high as 20.6% of total health costs in the USA, suggesting substantial room for cost savings through interventions that stem the cause of weight gain (Cawley and Meyerhoefer 2012; Finkelstein et al. 2009).[14]

Central explanations for the increase in American obesity are changes in food consumption or calorie expenditures. Prior work suggests that important factors in explaining the rise include the decreasing relative cost of food, the shift away from manual labor and to more sedentary work, increasing maternal labor supply, and the shift to a more sedentary lifestyle (Cawley 2011).[15] Additional work suggests that a peer's body weight in an individual's social network may influence their own body weight, suggesting obesity may be contagious (Christakis and Fowler 2007; Cunningham et al. 2012; Fletcher 2011). Last, other work suggests that technological change may be underlying these other proposed causes.[16]

Theoretically, as we describe in our Conceptual Framework section, the effect of Internet access on behavioral and environmental factors related to obesity remains ambiguous. Improved information on the negative health consequences of obesity, means to achieve a healthy weight, and access to social networks that promote healthy lifestyles could decrease obesity. At the same time, false information and access to social networks promoting negative health choices along with the potential for increased sedentary lifestyle suggest greater access may increase obesity. Last, increases in income could improve body weight measures if the income-health gradient is positive for body weight. This suggests that an empirical examination of the effects

[13] http://www.who.int/mediacentre/factsheets/fs311/en/ Accessed September 1, 2016.

[14] Some of these costs appear to be shifted to obese individuals. Bhattacharya and Bundorf (2009) find that obese individuals earn lower wages and that this serves to shift the cost of higher premiums onto the individual.

[15] A shift to a sedentary lifestyle is partially evidenced in the decreased availability of recreation spaces such as sidewalks and other open spaces.

[16] Lakdawalla et al. (2005) explore the role of welfare-improving technological change as underlying the drop in the relative price of food and the move to more sedentary occupations, and suggest that obesity is a side-effect of these technological changes.

could provide useful information regarding the overall effect of the increasing availability of broadband Internet on obesity.

4 Data

To proxy for Internet access, we use data on the broadband providers in a county over time. Data on broadband providers comes from the Federal Communication Commission's Form 477. This form documents the number of providers in each zip code in each year from 1999 to 2008, and the information is consolidated into a dataset available from the Federal Communications Commission. For the purposes of our analysis, we group zip codes to county and create population-weighted variables representing the fraction of the county with at least one broadband provider.[17] Although this is not a measure of individual use, it is correlated with use and serves as a good proxy for use (Guldi and Herbst 2017).

For data on health outcomes, we use the Behavioral Risk Factor Surveillance System (BRFSS) surveys from 1999 to 2007. The BRFSS is one of the largest data sets in the USA that provides information on adult health and health-related behaviors for a representative sample of non-institutionalized adults who are at least 18 years old.[18] Interviews are conducted by state health departments, assisted by the US Centers for Disease Control and Prevention, through monthly telephone interviews to collect data on health and health-related behaviors. The surveys consist of a set of standard core questions, optional modules, and state-specific questions.[19] We use the BRFSS to analyze the effects of county-level broadband availability on six outcomes covering weight and modifiable health behaviors that may affect weight and may change as a result of broadband availability. Our main weight measures are body mass index (BMI), an indicator for overweight status ($BMI \geq 25$), and an indicator for obese status ($BMI \geq 30$). In some specifications, we also look at extreme obesity ($BMI \geq 40$). Health behaviors include three indicator variables for any exercise activity in the last 30 days, any binge drinking events (five or more drinks in one occasion) in the last 30 days, and whether an individual currently smokes. Figure 2 demonstrates that the proportion of

[17] Data can be downloaded from http://transition.fcc.gov/wcb/iatd/comp.html. The documentation from the FCC indicates that these are "lists of geographical zip codes where service providers have reported providing high-speed service to at least one customer as of December 31, [of the relevant year]. No service provider has reported providing high-speed service in those zip codes not included in this list. An asterisk (*) indicates that there are one to three holding companies reporting service to at least one customer in the zip code. Otherwise, the list contains the number of holding companies reporting high-speed service. The information is from data reported to the FCC in Form 477."

[18] We would have liked to have examined childhood obesity, but the BRFSS does not survey individuals younger than 18. We chose to focus on the under 65 population since those age 65 and older are more likely be retired, and less likely to use the Internet in the same way as younger age groups; therefore, those who are 65 and older are likely to have a very different relationship between broadband introduction and health. Our analysis sample includes pregnant women, but our results are largely robust to excluding them from the sample.

[19] The core set of questions include a set of fixed core questions asked every year and a set of rotating core questions asked every other year. We focus on weight and health behavior outcomes from the fixed core of questions, but also utilize responses regarding the intensity of exercise that are part of the rotating core of questions in 2001, 2002, 2003, 2005, and 2007.

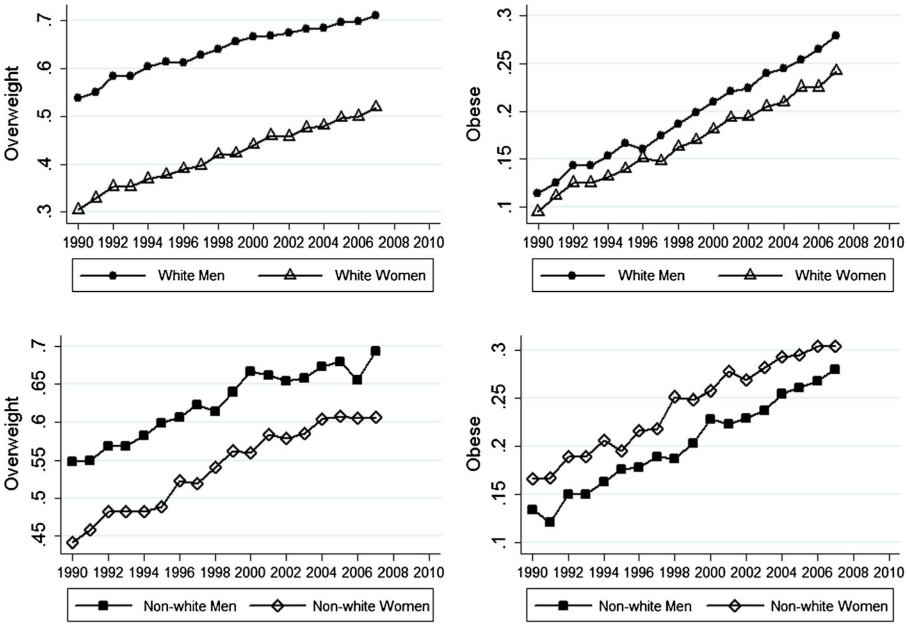

Fig. 1 Trends in fraction obese and overweight: 1990–2007. Source: author's calculations using the 1990–2007 Behavioral Risk Factor Surveillance System surveys

the population overweight rises over the same period as the proportion of the population with access to a broadband Internet provider.

Using the county geographic identifiers in the BRFSS, we match individuals to our county-level broadband availability measure in each year. We limit the sample to adults

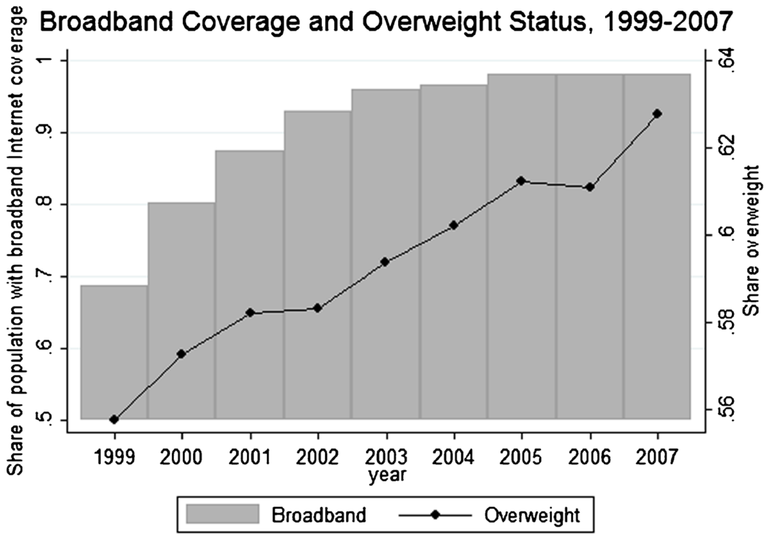

Fig. 2 Broadband coverage and share overweight: 1999–2007. Source: author's calculations using the 1999–2007 Behavioral Risk Factor Surveillance System surveys and the 1999–2007 Federal Communications Commission Form 447 data

age 18-64. We believe it is important to look at heterogeneity by race and gender due to cultural and behavioral differences in responses to Internet access and health interventions. To look at this heterogeneity, we further stratify the sample into white men, white women, non-white men, and non-white women. We drop observations from LA due to changes in infrastructure associated with Hurricane Katrina and observations from VA due to a large number of unmatched zip codes in the FCC data. Finally, we drop all observations without county identifiers and those missing any demographic control variables (age, gender, race, marital status, and education). This results in an unbalanced panel of counties. Because some counties are not consistently in the sample from 1999 to 2007, we restrict the sample to a balanced panel of counties as a robustness check and find that our baseline results do not substantially change (see Appendix Table 13).

Table 1 shows summary statistics for the variables used in the analysis. While the average county broadband availability increased from 68.9% in 1999 to 98.2% in 2007 in the USA as a whole, respondents in the BRFSS tend to be in counties with higher levels of broadband coverage with an average county-level broadband coverage of 95.6% in 1999 and 99.5% in 2007. The average age in the sample is approximately 40 years old, 67.7% are white, 58.7% are married, and 34.1% have at least a bachelor's degree.

We augment our BRFSS data with additional county-level covariates including the unemployment rate and real gross domestic product per capita. Table 2 shows the means of socio-economic characteristics for US counties, the subset of counties in our sample ("BRFSS counties"), and the subset of counties not in our sample ("non-BRFSS counties"). Overall, the BRFSS counties in our sample are wealthier, more urban, more educated, and spend more on social welfare than the counties not in our sample. The counties in the BRFSS, however, represent 93.9% of the US population, which suggests that the omitted counties are largely sparsely populated and our estimates represent the majority of individuals living in the USA.[20]

5 Methods

We use within-county changes in broadband providers to identify the impact of Internet access on health outcomes in adults. We perform the analysis separately by race and gender for several reasons. First, as shown in the raw data (Fig. 1), for both levels and trends there are clear differences by race in the rates of obesity and overweight by race and gender. Second, as we mention in the conceptual framework section, during our period of study, women were more likely than men to seek out health information via the Internet. This suggests that there may be important differences by gender in the behavioral health responses to Internet access. Third, adult interactions with the health system vary by race and gender. For example, data collected by the Center for Disease Control during 1997 to 1998, contemporaneous with the beginning of our study period, demonstrates that the number of doctor visits varies by gender and race.[21] Lastly, when

[20] Calculations made using Census county population estimates for 2000.

[21] See, for example, Utilization of Ambulatory Medical Care by Women: United States, 1997-1998. Vital Statistics and Health Series Report 13, No. 149. 51 pp. (PHS) 2001-1720.

Table 1 Summary statistics for full sample, 1999–2007

	Observations	Mean	Standard deviation	Min	Max
Demographics					
Age	1,416,133	39.67	12.69	18	64
Female	1,416,133	0.490		0	1
Less than high school degree	1,416,133	0.099			
High school graduate	1,416,133	0.275		0	1
Some college	1,416,133	0.280		0	1
Bachelor's degree or higher	1,416,133	0.346		0	1
White	1,416,133	0.684		0	1
Black	1,416,133	0.109		0	1
Hispanic	1,416,133	0.146		0	1
Other race	1,416,133	0.062		0	1
Married	1,416,133	0.587		0	1
County-level covariates					
Broadband coverage	1,416,133	0.989	0.043	0	1
Unemployment rate	1,416,133	4.99	1.57	0.7	29.7
Real per capita income ($)	1,416,133	35,430	10,435	13,319	167,901
Weight					
BMI	1,416,133	26.85	5.66	4.78	99.98
Overweight (BMI ≥ 25)	1,416,133	0.589		0	1
Obese (BMI ≥ 30)	1,416,133	0.229		0	1
Behaviors					
Any exercise in last 30 days	1,340,457	0.782		0	1
Binge drinking event in last 30 days	1,406,907	0.628		0	1
Currently smokes	1,412,533	0.233		0	1

BRFSS sampling weights used. Sample restricted to adults age 18–64. Observations from LA omitted due to changes in infrastructure related to Hurricane Katrina. Observations from VA omitted due to an unusually large number of unmatched zip codes in FCC data

attempting to lose weight, clinical evidence suggests different weight loss strategies are more effective for different gender and race groups (Jerome et al. 2015). Due to these observable differences by race and gender in the health outcomes we consider, the use of the Internet to seek out health information, and the utilization of healthcare, we expect the effects of Internet access on body weight to vary on these dimensions as well. Specifically, for each race by gender subgroup, we estimate the following reduced form model:

$$Y_{\text{icmt}} = \beta + \beta_1 \text{Broadband}_{\text{ct}} + \beta_2 X_{\text{icmt}} + \gamma_{\text{c}} + \lambda_{\text{m}} + \tau_{\text{t}} + \delta_{\text{ct}} + \epsilon_{\text{icmt}} \tag{1}$$

Here i indexes individuals, c indexes counties, m indexes months, and t indexes years. $\text{Broadband}_{\text{ct}}$ is the percentage of zip codes in a county with at least one broadband provider, and β_1 is the coefficient of interest. County, month of the year,

Table 2 County-level socio-economic variables

	All	BRFSS	Non-BRFSS	*p*-value
Per capita income ($)	23,916.26	24,603.85	22,269.36	0.000
N	3,110	2,194	916	
Unemployment rate	4.38	4.35	4.44	0.174
N	3,139	2,195	944	
Urban	31.7	34.7	18.0	0.000
N	2,606	2,129	477	
Total Population	288,764,448	271,189,824	17,574,624	
Mean Population	91,934	123,605	18,558	0.000
White (%)	84.4	85.1	82.9	0.001
Black (%)	8.9	8.4	10.2	0.002
Asian (%)	1.2	1.3	0.9	0.000
Other race (%)	2.9	2.9	3.0	0.973
N	3,141	2,194	947	
Education				
Less than a high school diploma (%)	22.6	21.7	24.7	0.000
High school diploma (%)	34.7	34.4	35.2	0.002
Some college or associate's degree (%)	26.2	26.5	25.5	0.000
Bachelor's degree or higher (%)	16.5	17.4	14.6	0.000
N:	3,141	2,194	947	
Social welfare expenditures ($1,000s)				
Medicaid	70,455.63	93,884.80	14,338.16	0.000
N	3,110	2,194	916	
Social security income	10,350.85	13,550.34	2,244.99	0.000
N	3,060	2,194	866	
Earned income tax credit	9,801.14	12,831.27	2,495.53	0.000
N	3,104	2,194	910	
SNAP	4,833.13	6,152.47	1,310.62	0.000
N	3,013	2,192	921	

BRFSS counties are counties which appear in our BRFSS sample after being matched to our FCC data. Non-BRFSS counties are counties which do not appear in our BRFSS sample. All data is for the year 2000 except for the indicator for urban which is from 1993. The reported *p*-value is from a two-sided difference-in-means test between counties that appear in our sample and counties that do not. Per capita income and social welfare expenditure data from the US Bureau of Economic Analysis. Unemployment rate is the average unemployment rate in 2000. Population, race, and education data from the US 2000 Decennial Census. Education data is for county population age 25 and older. Urban is defined according to the 1993 Rural-Urban continuum codes provided by the US Department of Agriculture.

and year fixed effects are represented by γ_c, λ_m, and τ_t, respectively. Including these fixed effects absorbs time invariant differences in health across counties, national differences in health specific to months of the year, and national differences in health across years. δ_{ct} is a county-specific linear time trend, which we include in many of our specifications so that identification comes from deviations from trends within counties and over time. Y_{icmt} is either BMI, an indicator for overweight, an indicator for obese, or one of our other adult health outcomes of interest.[22]

In Eq. 1, β_1 is identified from within county changes in health that coincide with within county changes in coverage, holding national average health within a year constant. A key assumption behind the identification of β_1 is that there are no trends in health prior to the entrance of broadband providers into counties relative to those who have not yet had an entrant. In other words, this model assumes that adult health was not improving (or declining) before a county experienced a change in broadband availability (relative to counties where at that time there was no change in providers).

To examine the validity of our assumption that the timing and degree of county broadband expansion is exogenous to county adult health, the differential trends assumption, we first turn to the careful work of other researchers who have used the same policy instrument, some of whom have examined pre-treatment trends directly for other outcomes (Atasoy 2013; Bellou 2015; Guldi and Herbst 2017; and Kolko 2012). Second, we directly test the robustness of our results to county-specific linear time trends for all of our estimates. We also test whether future broadband adoption influences past adult weight outcomes by including both leads and lags (and find that it does not). Finally, for counties we can observe in the years leading up to the rollout, we perform an analysis where we assign a "zero" for broadband in years before 1999 and find similar effects of broadband on body weight even once we add county trends. Unfortunately, due to data limitations we are unable to do a graphical analysis showing differences in treated versus untreated counties before the expansion of broadband with the data we have. While our measure of broadband availability begins in 1999, broadband providers are present to some degree in the majority of counties observable in the BRFSS data in 1999, making it difficult to distinguish the exact starting date of treatment in these counties. Relatedly, because our broadband Internet access proxy is a continuous measure of penetration, it is unclear how to define a discrete time in which broadband entrance into a county was initiated. Compounding these issues, not all BRFSS counties are observable in the pre-1999 era; and many that are available are not continuously in the data during the pre-1999 years due to the balance issues discussed in the Data section.[23] Even with this data limitation, our other, non-graphical tests, offer support of our parallel trends assumption.

Causal identification of β_1 also assumes that there are no contemporaneous unobserved changes in county policies, demographic composition, or characteristics that jointly induce broadband entrance and directly impact adult health. This assumption would be violated if, for example, county level policies designed to improve health also led to (or coincided with) the entrance of broadband providers. It is impossible to be entirely sure that this assumption holds, though we test it as rigorously as possible. Specifically, we add relevant time-varying county and individual-level observable characteristics (represented by X_{icmt}), to see if our estimates are sensitive to controlling for variables that would likely be correlated with unobservable changes. Individual-level controls include indicators for single year of age, education (high school, some

[22] We have also run models with month by year fixed effects (rather than separately controlling for year fixed effects and month fixed effects), which produced similar results to our baseline specification. These results are in Appendix Tables 14 and 15.

[23] This makes it difficult to do an event study which is ideally done with a balanced panel so as not to pick up compositional changes as counties enter and exit the sample in the graph.

college, and 4+ years of college), and marital status. In the non-white samples, we include an indicator for Hispanic and an indicator for all other reported races/ethnicities, using non-Hispanic black as our base group; and we control for economic conditions by including the county unemployment rate and real per capita income. In addition, we reviewed the literature on obesity prevention programs and related interventions that were plausibly changing during the time and found no reason to think these would be correlated with broadband penetration or cause a spurious relationship between broadband penetration and weight gain.[24] Finally, we note that any obesity prevention programs we are unaware of whose effectiveness is facilitated by Internet access will place a downward bias on our results. Therefore, even if this bias is a concern, our estimates would still rule out a zero or negative effect of broadband on weight gain in white women.

6 Results

6.1 Body weight

Table 3 shows our core results from the regressions of weight-related outcomes for the samples of white men and white women. Across the different outcomes, for white men the sign on our measure of broadband access are positive, but relatively small in magnitude and statistically insignificant. Overall, this suggests no consistent effect of Internet availability on health for white men. On the other hand, an increase in Internet availability has robust and moderately sized effects on a variety of measures of body weight for white women.[25] Starting at column 5, for white women a 10 percentage point increase in the fraction of the population in a county with at least one Internet provider would increase BMI by 0.1026 and the probability of being obese by 0.006. These represent effects that are 0.39 and 3.00% of the mean, respectively.

Moving across the columns of Table 3 shows that the results are generally robust to a variety of alternate specifications such as adding demographic and county level controls. One concern in difference-in-differences type models is that there may be differential trends in health between counties that expand broadband access and those that do not. However, as can be seen in moving from column 7 to 8 of Table 3, there is

[24] The following obesity-related interventions were prominent: food and cigarette taxes/prices, state mandatory physical education, nutrition and calorie labeling, and advertising of bad health behaviors (Cawley and Ruhm, 2011: pgs 97-109; Cawley 2015: pgs 256-258). For all of these policies, we generally did not see any reason to suggest that they would be correlated with broadband introduction. In addition, for many of these interventions, it was unclear how effective they were at changing obesity and nutrition. The recent research suggests that food taxes do not have an impact on obesity or nutrition (Cawley and Ruhm, 2011: pg 168), and there has been mixed evidence that cigarette taxes affect obesity and weight gain. Specifically, the effect of cigarette taxes and prices on increasing obesity is sensitive to specifications (such as how time is modeled) and could actually decrease obesity when dynamic effects are allowed for (Courtemanche 2009). Likewise, securing causal estimates of the effect of mandatory physical education on obesity has been difficult due to their likely being policy endogeneity biasing those estimates (Cawley and Ruhm, 2011: pg, 175). We consider advertising online of bad health behaviors to be a credible pathway for our effects. However, well-identified evidence on advertising is hard to come by and the literature that does exist shows mixed and often inconclusive results of advertising on risky health behavior (Cawley and Ruhm, 2011: pg 39)

[25] This is consistent with women engaging with online health information to a greater degree than men, as we mention in the Conceptual Framework section.

Table 3 Estimates of the effect of broadband availability on weight, white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.244	0.358	0.227	−0.205	1.026***	0.950**	0.853**	1.006**
	(0.268)	(0.250)	(0.249)	(0.348)	(0.365)	(0.358)	(0.353)	(0.461)
Mean		27.3				25.9		
R^2	0.025	0.079	0.079	0.085	0.029	0.082	0.082	0.086
Overweight (BMI ≥ 25)	0.016	0.032	0.029	0.010	0.069**	0.064**	0.067**	0.065*
	(0.027)	(0.025)	(0.026)	(0.038)	(0.028)	(0.028)	(0.027)	(0.034)
Mean		0.678				0.462		
R^2	0.018	0.082	0.082	0.087	0.024	0.074	0.074	0.077
Obese (BMI ≥ 30)	0.020	0.024	0.009	−0.024	0.061***	0.058***	0.048**	0.035
	(0.024)	(0.024)	(0.024)	(0.034)	(0.022)	(0.021)	(0.021)	(0.028)
Mean		0.233				0.198		
R^2	0.021	0.042	0.042	0.047	0.022	0.050	0.050	0.050
Extremely obese (BMI ≥ 40)	−0.001	−0.002	−0.001	−0.009	0.021**	0.020**	0.018*	0.022*
	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.010)	(0.012)
Mean		0.021				0.031		
R^2	0.015	0.019	0.019	0.026	0.012	0.021	0.021	0.025
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	474,723	474,723	474,723	474,723	651,627	651,627	651,627	651,627

Standard errors clustered at the county level. All regressions include month, year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, marital status, education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights.

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

little change in the coefficients after adding county linear trends.[26] Overall, we consider these estimates to be reasonably robust. We also estimated similar models for the non-white samples of men and women. These results, reported in Table 4, follow a similar pattern, though most coefficients are not statistically significant.[27] We also explored the estimates by age and found that the effects for white women are strongest for the youngest group (age 18 to 34), the ages with the highest rates of reported Internet use during this period.[28]

We explored the robustness of the body weight estimates in two additional analyses that involved extending our pre-period beyond the time for which we have broadband data available. First, we expanded our sample to also include years back to 1990. Since broadband Internet providers were virtually non-existent at this time, we set our county broadband penetration measure to zero for years 1990 to 1993 and omitted years 1994 to 1996, when broadband rollout started. This allows us to more convincingly test for robustness to county-linear trends by extending a pre-period. We show that estimated effects on body weight (see Appendix Table 16) are similar to our baseline estimates in Tables 3 and 4. Next, for white women only, we estimate a similar model but where we include the years 1994 to 1996 and assign all counties zero broadband coverage in these years (Appendix Table 17). We see that including these years leads to a slightly smaller coefficient estimate in column 2 (which includes 1994 to 1996) than in column 1 (which excludes 1994 to 1996), which is exactly what we expect due to measurement error since some of these counties are expanding broadband during this early period but are assigned zero broadband coverage. Finally, again for white women only, we estimate the effects separately for the 1999 to 2003 period, the time when the greatest rollout of broadband occurred, and for the 2004 to 2007 period, when broadband rollout was nearly completed. These estimates are reported in Appendix Table 18 and show that, indeed, the effects are concentrated in the earlier period. While Appendix Tables 17 and 18 show results for white women only, we ran these tests on the other demographic groups and found consistent results to our main estimates for these groups.

6.2 Health behaviors

If the Internet increases weight gain through increased sedentary activity, we would possibly see decreased reported exercise with Internet rollout. Similarly, exposure to lower quality health information, or the influence of expanded social networks, could increase obesity through worsening health behaviors. Though it is difficult to completely isolate any one of these mechanisms, in this section we explore changes in health behaviors as a pathway through which increased obesity occurs. During our sample period, the BRFSS consistently collects information on a number of health behaviors of interest: exercise, binge drinking, and smoking. The estimates of our model with these health behaviors as outcomes are in Table 5 (whites) and Table 6 (non-whites). Our

[26] While the coefficient on obesity loses statistical significance, the magnitude of the coefficient is qualitatively similar.

[27] These noisier effects on the non-white samples are potentially due to the smaller sample. An alternative explanation is that our broadband measure captures access less consistently for these group: though with somewhat larger effects on those who are affected.

[28] These estimates are available upon request.

Table 4 Estimates of the effect of broadband availability on weight, non-white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.878	0.596	0.251	0.166	1.493*	1.419*	1.216	−0.051
	(0.798)	(0.723)	(0.689)	(1.015)	(0.776)	(0.737)	(0.750)	(1.089)
Mean		27.33				27.37		
R^2	0.038	0.101	0.101	0.120	0.042	0.138	0.138	0.145
Overweight (BMI ≥ 25)	−0.033	−0.051	−0.092	−0.047	0.069	0.064	0.069	−0.038
	(0.073)	(0.065)	(0.067)	(0.102)	(0.069)	(0.068)	(0.068)	(0.082)
Mean		0.664				0.585		
R^2	0.026	0.104	0.104	0.115	0.032	0.127	0.127	0.135
Obese (BMI ≥ 30)	−0.015	−0.026	−0.039	−0.006	0.088	0.082	0.068	0.062
	(0.068)	(0.067)	(0.064)	(0.092)	(0.060)	(0.057)	(0.054)	(0.082)
Mean		0.241				0.277		
R^2	0.031	0.066	0.066	0.078	0.032	0.092	0.092	0.098
	0.052***	0.058***	0.051***	0.042	0.022	0.022	0.017	−0.005
	(0.018)	(0.018)	(0.018)	(0.030)	(0.025)	(0.024)	(0.027)	(0.024)
Mean		0.023				0.048		
R^2	0.027	0.033	0.033	0.050	0.022	0.037	0.037	0.044
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	111,669	111,669	111,669	111,669	178,114	178,114	178,114	178,114

Standard errors clustered at the county level. All regressions include month, year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, race (Hispanic, all other races/ethnicities; non-Hispanic black as base group), marital status, and education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 5 Estimates of the effect of broadband availability on health behaviors, white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any exercise in last 30 days	0.089**	0.099***	0.099***	0.070	0.053	0.058	0.065*	0.148***
	(0.039)	(0.038)	(0.039)	(0.058)	(0.038)	(0.037)	(0.037)	(0.057)
Mean		0.831				0.806		
R^2	0.028	0.078	0.078	0.084	0.031	0.078	0.078	0.082
N	447,706	447,706	447,706	447,706	617,678	617,678	617,678	617,678
Any binge drinking events in last 30 days	0.105***	0.084***	0.073**	0.035	0.122***	0.120***	0.106***	0.063
	(0.029)	(0.029)	(0.028)	(0.044)	(0.037)	(0.037)	(0.036)	(0.047)
Mean		0.619				0.575		
R^2	0.067	0.112	0.112	0.118	0.091	0.122	0.122	0.127
N	470,818	470,818	470,818	470,818	649,059	649,059	649,059	649,059
Currently smokes	0.017	−0.007	−0.012	−0.012	0.041*	0.037*	0.031	−0.000
	(0.029)	(0.026)	(0.027)	(0.033)	(0.023)	(0.022)	(0.022)	(0.030)
Mean		0.252				0.234		
R^2	0.024	0.116	0.116		0.022	0.117	0.117	0.118
N	473,559	473,559	473,559	473,559	650,073	650,073	650,073	650,073
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X

Note: Standard errors clustered at the county level. All regressions include month, year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, marital status, education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 6 Estimates of the effect of broadband availability on health behaviors, non-white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Any exercise in last 30 days	−0.108	−0.093	−0.094	−0.096	−0.020	−0.007	−0.027	0.023
	(0.084)	(0.082)	(0.085)	(0.068)	(0.065)	(0.066)	(0.071)	(0.103)
Mean		0.734				0.669		
R^2	0.028	0.087	0.087	0.100	0.023	0.062	0.062	0.071
N	105,574	105,574	105,574	105,574	169,435	169,435	169,435	169,435
Any binge drinking events in last 30 days	0.299***	0.293**	0.223**	0.282**	0.086	0.066	0.038	0.074
	(0.114)	(0.116)	(0.102)	(0.128)	(0.063)	(0.059)	(0.056)	(0.086)
Mean		0.684				0.708		
R^2	0.053	0.077	0.078	0.091	0.048	0.078	0.079	0.087
N	109,997	109,997	109,997	109,997	177,033	177,033	177,033	177,033
Currently smokes	−0.060	−0.091	−0.129*	−0.144*	0.025	0.014	0.021	0.090
	(0.081)	(0.077)	(0.075)	(0.087)	(0.056)	(0.053)	(0.055)	(0.060)
Mean		0.254				0.165		
R^2	0.028	0.071	0.071	0.084	0.040	0.077	0.077	0.087
N	111,259	111,259	111,259	111,259	177,642	177,642	177,642	177,642
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X

Standard errors clustered at the county level. All regressions include month, year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, race (Hispanic, all other races/ethnicities; non-Hispanic black omitted), marital status, and education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

estimates show that Internet access increases harmful health behaviors for white men and women. Specifically, we see that for white women, Internet access increases exercise, binge drinking, and smoking, although the estimated effects on binge drinking and smoking are not robust to including county linear trends.[29] For non-whites, however, except for binge drinking, our estimates do not provide statistically meaningful evidence on modifiable health behaviors.

Although our main estimates reveal that Internet access leads to weight gain, our Table 5 estimates also show that for white women broadband coverage increases exercise activity, which is generally thought to lower body weight. The exercise variable we use, however, does not capture exercise frequency or intensity. It may be that Internet access fails to improve exercise behaviors to a level of intensity that improves body weight. Another possible explanation for increasing exercise and weight gain is that white women who begin a new exercise regime in response to information about exercise on the Internet over-estimate how many calories they burn and in turn over-compensate with the calories they eat and drink. This is consistent with anecdotal evidence that some people gain weight when they start an exercise program and epidemiological work which attempts to explain why people who exercise lose less weight than expected (Miller et al. 1997; Thomas et al. 2012; Melanson et al. 2013; Dhurandhar et al. 2015).[30] Given the high calorie content of alcohol, increases in binge drinking are also consistent with increased weight in white women. While increases in exercise and binge drinking are present for white men (in models without county linear time trends), the estimated effect of exercise is larger than the effect of binge drinking, which may mean that any weight gain from drinking is overcome with the additional exercise.

To better understand the exercise results, we look at exercise intensity as an outcome. These results are in Table 7 (whites) and Table 8 (non-whites). Here the outcome variable is either an indicator for moderate exercise (as opposed to no exercise or vigorous exercise) or vigorous exercise (relative to no exercise or only moderate exercise). For white men, the standard errors are large making it difficult to draw a firm conclusion on how Internet changes exercise intensity. For white women, we see that for those who exercised, the broadband coverage increased moderate exercise and not vigorous exercise, which is consistent with a story of Internet access failing to improve health behaviors at a level of intensity that offsets the increased weight gain from drinking. However, a caveat to these results for white women is that they are not robust to including linear time trends. Estimated effects for non-white men and women are generally not statistically significant, except when we include linear time trends.

6.3 Role of income

We explore the relationship between income and broadband access by adding an interaction between our broadband variable and whether the observation was in a relatively high or low-income county (based on pre-period 1999 income levels). This

[29] Since smoking is an appetite suppressant, it can be associated with declines in weight. However, we take the increase in smoking as evidence for a more general story of broadband expansions causing overall worse health behaviors which in turn outweighs the benefits of decreased food consumption from smoking.

[30] See for example, https://www.huffingtonpost.com/kelly-coffey/a-trainer-comes-clean-abo_b_5977286.html, accessed March 28, 2018.

Table 7 Estimates of the effect of broadband availability on health behaviors, white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Only moderate activity in a usual week	0.068	0.078	0.092	0.078	0.138*	0.142*	0.148*	0.056
	(0.069)	(0.073)	(0.074)	(0.074)	(0.082)	(0.082)	(0.082)	(0.107)
Mean		0.285				0.436		
R^2	0.022	0.054	0.054	0.063	0.020	0.043	0.043	0.049
Only vigorous activity in a usual week	0.023	0.022	0.020	0.026	0.013	0.012	0.013	0.008
	(0.031)	(0.078)	(0.032)	(0.044)	(0.016)	(0.016)	(0.016)	(0.29)
Mean		0.036				0.014		
R^2	0.018	0.020	0.021	0.029	0.011	0.013	0.013	0.019
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	234,834	234,834	234,834	234,834	330,014	330,014	330,014	330,014

Standard errors clustered at the county level. All regressions include month, year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, marital status, education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 8 Estimates of the effect of broadband availability on health behaviors, non-white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Only moderate activity in a usual week	0.142	0.138	0.137	0.526**	−0.145	−0.128	−0.087	0.203
	(0.102)	(0.103)	(0.116)	(0.224)	(0.118)	(0.123)	(0.130)	(0.200)
Mean		0.276				0.432		
R^2	0.033	0.059	0.059	0.080	0.028	0.038	0.038	0.056
Only vigorous activity in a usual week	0.089	0.085	0.060	0.161**	0.093	0.091	0.098	0.137**
	(0.010)	(0.097)	(0.107)	(0.065)	(0.068)	(0.067)	(0.066)	(0.060)
Mean		0.069				0.031		
R^2	0.033	0.042	0.042	0.065	0.025	0.029	0.029	0.043
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	53,218	53,218	53,218	53,218	87,850	87,850	87,850	87,850

Standard errors clustered at the county level. All regressions include month, year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, race (Hispanic, all other races/ethnicities; non-Hispanic black as base group), marital status, and education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

allows for heterogeneous effects of broadband Internet access for those who are already living in affluent areas versus poorer areas.[31] The results in Table 9 are not statistically significant, although there is a pattern showing that for richer counties (as measured by either being above the median, or at the 80th percentile), broadband access increases weight gain and obesity; with a larger relative effect compared to lower income counties. Potentially, this relatively lower negative effect on health for low income counties is because lower income areas had more to gain from the economic activity associated with broadband. This may suggest that there are important differences in how income interacts with broadband, but income itself is not necessarily the principal channel explaining declines in health. We are hesitant to draw strong conclusions from these estimates, however, since they lack statistical significance.

6.4 Other potential mechanisms

In addition to health behaviors and income, we explore several other potential mechanisms. Using the BRFSS data, we were able to estimate how broadband Internet access affects the following outcomes (all indicator variables): (1) employed at the time of survey, (2) currently trying to lose weight, (3) has any type of health insurance, (4) had a routine check-up in the last 12 months, and (5) had a cholesterol check in the last 12 months. These results are in Appendix Table 19. Unfortunately, due to a lack of statistical precision, these estimates do not provide clear evidence on a single mechanism.

6.5 Falsification tests

As Fig. 2 shows, obesity rates were trending up since the 1990s and 2000s during the rollout of broadband Internet. While Figs. 1 and 2 show national trends, it is possible that there is a spurious pre-trend in body weight for counties that later increase their broadband Internet coverage. We conduct a falsification test by check the timing of broadband availability and changes in body weight by replacing our broadband measure with a three year lead of our broadband measure. More specifically, we estimate the following model for each of our three weight measures:

$$Y_{\mathrm{icmt}} = \beta + \beta_1 \mathrm{Broadband}_{\mathrm{ct+3}} + \beta_2 X_{\mathrm{icmt}} + \gamma_{\mathrm{c}} + \lambda_{\mathrm{m}} + \tau_{\mathrm{t}} + \delta_{\mathrm{ct}} + \epsilon_{\mathrm{icmt}} \quad (2)$$

For this analysis, we match individual-level data from the 1996–2004 BRFSS to their corresponding county-level broadband measure 3 years in the future.[32] We include the same control variables in X_{icmt}, fixed effects, and county-specific time trends as in Eq. 1. Results from these regressions are reported in Table 10 (whites) and Table 11 (non-whites). Our estimates are not statistically significant for any regression in either table. Furthermore, the estimates are of mixed signs and when the estimates are positive, the magnitudes are generally smaller than our main estimates (Tables 3 and 4).

[31] We use 1999 county income as a proxy that is correlated with yearly county income but that is not directly affected by broadband rollout.

[32] For example, individual observations from Middlesex County, MA in the 1996, BRFSS are matched to the 1999 broadband measure for Middlesex County, MA.

Table 9 Effect of broadband on weight interacted with county income indicators in 1999, white samples

	White men						White women					
	BMI		Overweight		Obese		BMI		Overweight		Obese	
Broadband	−0.373 (0.461)	−0.168 (0.411)	−0.048 (0.043)	−0.013 (0.041)	0.006 (0.032)	−0.025 (0.043)	0.347 (0.689)	0.732 (0.436)	0.015 (0.032)	0.022 (0.043)	−0.004 (0.038)	0.041 (0.035)
Broadband × (above median county income in 1999)	0.286 (0.620)		0.100 (0.061)		−0.051 (0.064)		1.140 (0.919)		0.086 (0.064)		0.068 (0.052)	
Broadband × (80th percentile county income in 1999)		0.382 (0.744)		0.105 (0.078)		0.034 (0.074)		0.923 (0.978)		0.121 (0.078)		0.009 (0.063)
Broadband × (20th percentile county income in 1999)		−1.408 (0.952)		−0.068 (0.075)		−0.081 (0.074)		0.082 (1.082)		0.066 (0.071)		-0.065 (0.066)
N	474,723	474,723	474,723	474,723	474,723	474,723	651,627	651,627	651,627	651,627	651,627	651,627

Standard errors clustered at the county level. All regressions include month, year, and county fixed effects, demographic controls (indicator variables for age, marital status, and education (high school graduate, some college, and bachelor's degree or higher)), county unemployment rate and real per capita income, and county-specific linear time trends. "Broadband × (80th percentile county income in 1999)" is an interaction between our broadband measure and an indicator for whether county is in the 80th percentile of county real per capita income in 1999. "Broadband × (20th percentile county income in 1999)" is an interaction between our broadband measure and an indicator for whether county is in the 20th percentile of county real per capita income in 1999. "Broadband × (above median county income in 1999)" is an interaction between our broadband measure and whether county is above median county real per capita income in 1999. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 10 Estimates of the effect of future broadband availability ($t + 3$) on weight, white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.390	0.399	0.305	−0.101	−0.096	−0.058	−0.198	−0.668
	(0.306)	(0.298)	(0.313)	(0.428)	(0.358)	(0.339)	(0.339)	(0.451)
R^2	0.020	0.079	0.079	0.083	0.025	0.082	0.082	0.086
Overweight (BMI ≥ 25)	0.024	0.027	0.033	−0.032	−0.017	−0.013	−0.019	−0.014
	(0.037)	(0.036)	(0037)	(0.046)	(0.030)	(0.029)	(0.030)	(0.043)
R^2	0.016	0.080	0.080	0.084	0.020	0.073	0.073	0.078
Obese (BMI ≥ 30)	0.026	0.026	0.022	0.011	0.019	0.019	0.008	−0.026
	(0.026)	(0.026)	(0.026)	(0.036)	(0.025)	(0.025)	(0.024)	(0.032)
R^2	0.015	0.038	0.038	0.042	0.017	0.044	0.044	0.048
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	340,242	340,242	340,242	340,242	440,138	440,138	440,138	440,138

Standard errors clustered at the county level. We define "future broadband availability" as the level of broadband availability in a county three years from each year of the BRFSS sample, beginning in 1996. All regressions include month, year, and county fixed effects. Columns 2 and 5 add demographic controls: indicator variables for age, marital status, education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 6 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 11 Estimates of the effect of future broadband availability ($t + 3$) on weight, non-white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	1.608	1.309	0.600	0.236	0.710	0.675	0.256	−0.650
	(1.113)	(0.974)	(0.977)	(1.150)	(0.742)	(0.660)	(0.697)	(1.010)
R^2	0.037	0.103	0.104	0.119	0.040	0.142	0.143	0.149
Overweight (BMI ≥ 25)	0.194	0.170	0.105	0.091	0.071	0.072	0.041	−0.043
	(0.121)	(0.107)	(0.109)	(0.113)	(0.076)	(0.072)	(0.071)	(0.084)
R^2	0.024	0.100	0.100	0.11	0.031	0.130	0.130	0.137
Obese (BMI ≥ 30)	0.024	0.009	−0.026	−0.058	0.008	0.001	−0.016	−0.100
	(0.065)	(0.059)	(0.061)	(0.087)	(0.059)	(0.059)	(0.062)	(0.099)
R^2	0.026	0.058	0.058	0.069	0.029	0.086	0.086	0.093
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	81,445	81,445	81,445	81,445	120,588	120,588	120,588	120,588

Note: Standard errors clustered at the county level. We define "future broadband availability" as the level of broadband availability in a county three years from each year of the BRFSS sample, beginning in 1996. All regressions include month, year, and county fixed effects. Columns 2 and 5 add demographic controls: indicator variables for age, race (Hispanic, all other races/ethnicities; non-Hispanic black as base group), marital status, and education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 6 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Next, we examine this in a slightly different way by estimating a model that includes a set of 3 lags and 3 leads of the coverage variable as follows:

$$Y_{\text{icmt}} = \beta + \sum_{n=-3}^{3} \rho_n \text{Broadband}_{\text{ct}+n} + \beta_3 X_{\text{icmt}} + \gamma_{\text{c}} + \lambda_{\text{m}} + \tau_{\text{t}} + \delta_{\text{ct}} + \epsilon_{\text{icmt}} \quad (3)$$

Here, $\sum_{n=-3}^{3} \rho_n \text{Broadband}_{\text{ct}+n}$: reflects the current period, as well as 1, 2, and 3 years before and after the current period. This model tests whether the current change in weight is occurring with the current change in broadband or with a change in broadband in earlier or later years. Because $\text{Broadband}_{\text{ct}+n}$, $n \in [-3..3]$, are not dichotomous variables, there is not a reference/excluded group. In all of our models, effects are identified off of the relative changes in levels of penetration relative to changes in levels of obesity across both counties and years. If there were pre-trends in the model, then we would expect current changes in weight to be correlated with future levels of broadband. The results of the model for white women are reported in Table 12.[33] Across these three models, the coefficient on current broadband ($\text{Broadband}_{\text{ct}}$) is similar in magnitude to our main estimates in Table 3 and statistically significant for BMI and overweight. Furthermore, the estimates for broadband leads and lags are not statistically significant and small relative to our main estimates.

Taken together, these tests suggest that our results are not driven by a spurious correlation between broadband availability and body weight.

6.6 Multiple hypothesis testing

Another concern is that with so many outcomes tested, we might have estimated coefficients that are, by chance, statistically significant. To address the issue of multiple hypothesis testing (MHT), we use a reporting index test akin to those implemented by Anderson (2008) and Kling et al. (2007).[34] When dealing with multiple outcomes, the procedure to construct an index is to (1) separate outcomes into "families" of similar outcomes, (2) normalize the outcomes for "sign agreement" such that all beneficial (harmful) outcomes have a positive (negative) sign, and (3) construct a weighted average through demeaning and dividing by the control group standard deviation. This addresses MHT concerns because for each family of outcomes there is now only a single hypothesis test. Below we discuss MHT within the context of subgroups, weight-related outcomes, and health health-behavior related outcomes.

First, while we would like to do a similar correction for subgroups, the current applied microeconometric literature has only developed subgroup MHT for the experimental setting (List et al. 2016). One problem with applying a correction (such as an index) to subgroups is that it would implicitly assume that the reactions across subgroups should be the same, something we believe is unlikely to be true since there are known health and health behavior differences by gender and race, which we describe in the Conceptual

[33] We also estimate models for the other race-gender groups and find the current period estimates are congruent with the main paper table estimates (no statistically significant effect of current broadband). We do not report them in the main paper but are happy to share the estimates upon request.

[34] The most straightforward approach is the Bonferroni correction, but it is viewed as overly conservative (Christensen and Miguel, 2016; Ross et al., 2008)

Table 12 Falsification test using 3-year leads and lags of broadband, white women

	Body mass index (BMI)	Overweight (BMI ≥ 25)	Obese (BMI ≥ 30)
Broadband_{ct} (current broadband)	0.728*	0.054*	0.035
	(0.372)	(0.032)	(0.024)
Broadband_{ct+3}	−0.056	0.000	−0.000
	(0.536)	(0.052)	(0.041)
Broadband_{ct+2}	0.504	0.006	0.029
	(0.422)	(0.037)	(0.034)
Broadband_{ct+1}	0.331	0.036	0.039
	(0.387)	(0.038)	(0.026)
Broadband_{ct-1}	0.022	−0.018	−0.009
	(0.309)	(0.028)	(0.022)
Broadband_{ct-2}	−0.239	0.004	0.007
	(0.352)	(0.025)	(0.022)
Broadband_{ct-3}	−0.069	0.004	−0.011
	(0.297)	(0.023)	(0.021)
N	751,520	751,520	751,520

Standard errors clustered at the county level. We include our measure of county broadband coverage in the current year (**Broadband_{ct}**) and broadband measures with three year leads and lags (**$\text{Broadband}_{ct\pm n}$**, where **$n = 1,2,3$**). All regressions include month, year, and county fixed effects, demographic controls (indicator variables for age, marital status, and education (high school graduate, some college, and bachelor's degree or higher)), county unemployment rate and real per capita income, and county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Framework and Background sections. Along these lines, we believe the best evidence against concerns about MHT is that we see statistically significant effects for white women: the group that is the most likely to use the Internet to access health information and therefore the group who arguably has the highest propensity to be treated.

Second, while we could apply such an index to our "family" of weight-related outcomes (BMI, overweight, obese, and extremely obese), each of these are (essentially) already different transformations of the underlying measure of BMI. Therefore, it seems to make more sense to simply use BMI as our "summary index" in this case. As reported in Table 3, BMI is statistically significant for white women across all our models. Likewise, BMI follows the same patterns as the other weight outcomes for white women as well as the other subgroups we examine as shown in Tables 3 and 4, suggesting that BMI is a good summary measure of the weight outcomes we explore in the paper.

Last, we consider the other health outcomes we examine: exercise, drinking, and smoking. We create a single "risky health behavior" index using these variables to address MHT issues. We normalize these variables to have a mean zero and standard deviation of one in the overall sample, such that a large value for these variables represents worse health. Then we create the index by averaging these three composite variables together into a single index of negative health behaviors such that a positive coefficient reflects more drinking and smoking and less exercise. These results are reported in Appendix Table 20 and show similar inference when using the health behavior summary index of negative health behaviors as with models with each of the individually measured health outcomes where we also control for county linear trends (see Tables 5 and 6, columns 4 and 8), though we note that for white women the index measure, while positive, is no longer statistically significant and that for non-white women is now statistically significant. Taken together, our results suggest that while we provide strong evidence that broadband coverage leads to increases in white women's body weight, we are unable to unequivocally demonstrate that this happens through changes in an index of health behaviors. However, when creating our index we have assumed that *lower* exercise is a *worse* health behavior. Yet, as we discuss above within the context of our main results, some anecdotal and epidemiological evidence suggests women *gain* weight with *increased* exercise. Therefore, it may be that our exercise estimates are, in part, responsible for the weight gains we observe.

7 Conclusion

Obesity has become an increasingly important public health issue in the USA and many other countries. Hypothesized causes for this uptick include declining relative cost of food and decreasing share of the population working in labor-intensive occupations. In this paper, we hypothesize that the Internet may also influence the obesity rate.

We use the rollout of broadband Internet providers as a plausible source of exogenous variation in Internet use to identify the effects of Internet use on obesity and body weight. We show that greater broadband coverage increases body weight and has both positive and negative effects on modifiable adult health behaviors including exercise, smoking, and drinking. A 10% increase in broadband availability increases obesity among white women by 0.0035, which represents an effect of 1.8% of the mean. Referring to Fig. 1, between 1990 and 2007, obesity increased by 15% for white women; suggesting that while the Internet is by no means the driving force behind this increase, it is at least part of the story. Our back of the envelope calculation suggests that increased medical costs for white women from obesity due to broadband Internet coverage over this time came to approximately \$2.2 billion.[35] On the other hand, our results show that broadband coverage was not a substantial cause of increased body weight for non-white women or for men over this time. This discrepancy is likely due to differences in behavioral responses to broadband Internet availability and differences in how Internet use influences health behaviors.

How can we explain the mechanisms behind our findings that broadband Internet availability increases obesity in white women? Theoretically, if increases in broadband availability improve information, this should lead to more optimal consumer choices. However, as we show, such choices do not necessarily mean health improves: greater networking opportunities available through the Internet may result in peers having greater influence over positive or negative health behaviors. Indeed, a number of papers have linked the Internet to expanding social circles (Wellman et al. 1996; Wellman and Gulia 1999; Wellman et al. 2001; Zhao 2006). While a pure information effect should decrease the likelihood of obesity, peer effects on health behaviors may work in either direction. Our estimates suggest broadband Internet availability increases drinking for men and women.

[35] We calculated this by multiplying our estimate of the effect of increasing obesity for white women (a 3.5 percentage point increase in obesity) by the change Internet providers over the years of our sample (a 29.3% increase) by the population of adult white women in the USA in 2005 (68,013,866). This suggests that the Internet pushed 1.2 million white women into obesity. According to Cawley and Meyerhoefer (2012), annual cost estimates for obesity are \$3613 (women) or \$2739 (white), suggesting an increase in costs of approximately \$2.2 billion.

While Internet use is a sedentary activity, counter to this explanation, we find some evidence of increases in exercise for men and women. While it is possible that increased exercise can lead to weight gain as appears to be occurring for white women, the large effects on exercise for white men appear to cancel out any weight increase. On the other hand, there are relatively small effects of exercise for white women. Taken together, broadband Internet coverage appears to provide both positive and negative health benefits with a net effect of increasing obesity rates for white women.

Acknowledgements We owe special thanks to Ken Couch, David Slutsky, Lyudmyla Sonchak, and to participants of the 2018 American Economic Association Annual Meeting, 2017 World Congress of the International Health Economics Association, 2017 Eastern Economic Association Annual Meeting, and 2016 Southern Economic Association Annual Meeting and two anonymous reviewers for their helpful comments on this paper.

Funding information The authors received no funding for this study.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Table 13 Effect of broadband on weight using "balanced counties" samples

	Men		Women	
	White	Non-white	White	Non-white
BMI	−0.251	−0.322	1.139**	0.239
	(0.377)	(1.873)	(0.532)	(2.022)
Mean	27.23	27.26	25.68	27.29
Overweight (BMI ≥ 25)	0.012	−0.143	0.078**	−0.064
	(0.041)	(0.193)	(0.040)	(0.147)
Mean	0.670	0.659	0.448	0.581
Obese (BMI ≥ 30)	−0.031	−0.006	0.052*	0.103
	(0.038)	(0.155)	(0.031)	(0.144)
Mean	0.224	0.237	0.188	0.272
Extremely obese (BMI ≥ 40)	−0.009	0.041	0.023*	−0.011
	(0.011)	(0.50)	(0.014)	(0.042)
Mean	0.019	0.022	0.029	0.046
N	365,921	85,766	496,779	136,261

Standard errors clustered at the county level. Each sample only includes observations from counties that are in the BRFSS every year in our sample period. All regressions include indicator variables for age, marital status, education (high school graduate, some college, and bachelor's degree or higher), county-level unemployment rate and county-level real per capita income, month fixed effects, year fixed effects, county fixed effects, and county-specific linear time trends. Regressions for the non-white samples also include indicators for Hispanic and other race/ethnicities (non-Hispanic blacks as base group). All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 14 Estimates of the effect of broadband availability on weight using month by year fixed effects specifications, white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.230	0.343	0.212	−0.217	1.029***	0.954***	0.855**	1.002**
	(0.267)	(0.250)	(0.249)	(0.348)	(0.364)	(0.356)	(0.351)	(0.460)
Mean		27.34				25.85		
Overweight ($BMI \geq 25$)	0.016	0.032	0.028	0.010	0.068**	0.063**	0.066**	0.064*
	(0.027)	(0.025)	(0.026)	(0.038)	(0.028)	(0.028)	(0.027)	(0.034)
Mean		0.678				0.462		
Obese ($BMI \geq 30$)	0.019	0.023	0.009	−0.024	0.062***	0.058***	0.048**	0.035
	(0.024)	(0.024)	(0.024)	(0.034)	(0.021)	(0.021)	(0.021)	(0.027)
Mean		0.233				0.198		
Extremely obese ($BMI \geq 40$)	−0.001	−0.002	−0.001	−0.009	0.021**	0.020**	0.018*	0.022*
	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)	(0.009)	(0.010)	(0.012)
Mean		0.021				0.031		
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	474,723	474,723	474,723	474,723	651,627	651,627	651,627	651,627

Standard errors clustered at the county level. All regressions include month by year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, marital status, education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 15 Estimates of the effect of broadband availability on weight using month by year fixed effects specifications, non-white samples

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.839	0.538	0.204	0.154	1.531**	1.440**	1.245*	−0.083
	(0.785)	(0.712)	(0.680)	(1.013)	(0.772)	(0.731)	(0.744)	(1.081)
Mean		27.33				27.37		
Overweight (BMI ≥ 25)	−0.030	−0.050	−0.091	−0.048	0.071	0.064	0.069	−0.041
	(0.073)	(0.065)	(0.067)	(0.099)	(0.068)	(0.068)	(0.068)	(0.082)
Mean		0.664				0.585		
Obese (BMI ≥ 30)	−0.019	−0.032	−0.046	−0.006	0.092	0.085	0.071	0.063
	(0.067)	(0.067)	(0.064)	(0.092)	(0.059)	(0.056)	(0.053)	(0.081)
Mean:		0.241				0.277		
Extremely obese (BMI ≥ 40)	0.050***	0.046***	0.049***	0.041	0.023	0.022	0.018	−0.005
	(0.018)	(0.018)	(0.018)	(0.029)	(0.025)	(0.024)	(0.027)	(0.024)
Mean		0.023				0.048		
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	111,669	111,669	111,669	111,669	178,114	178,114	178,114	178,114

Standard errors clustered at the county level. All regressions include month by year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, race (Hispanic, other race/ethnicities; non-Hispanic black as base group), marital status, education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 16 Robustness test, 1990–2007 BRFSS with broadband set to zero before 1999, excludes 1994–1996

	White men	White women	Non-white men	Non-white women
BMI	0.239	0.915**	0.003	0.603
	(0.319)	(0.436)	(0.786)	(0.888)
Mean	27.07	25.51	27.11	27.09
R^2	0.097	0.099	0.128	0.157
Overweight (BMI ≥ 25)	0.046	0.067*	−0.048	−0.018
	(0.034)	(0.034)	(0.082)	(0.078)
Mean	0.656	0.435	0.649	0.567
R^2	0.094	0.089	0.121	0.144
Obese (BMI ≥ 30)	0.006	0.035	−0.0381	0.049
	(0.029)	(0.023)	(0.069)	(0.069)
Mean	0.214	0.181	0.226	0.263
R^2	0.053	0.058	0.082	0.102
N	519,614	706,911	122,207	192,670

Standard errors clustered at the county level. All regressions control for age, marital status, education (high school graduate, some college, and bachelor's degree or higher), average yearly county unemployment rate, county per capita income, county-specific linear time trends, and include month, year, and county fixed effects. Regressions for the non-white samples also include indicators for Hispanic and other race/ethnicities (non-Hispanic black as base group). All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 17 Robustness test including pre-broadband era BRFSS samples, white women

	Add 1990–1993 BRFSS (1)	Add 1990–1996 BRFSS (2)
BMI	0.915**	0.832**
	(0.436)	(0.390)
Mean	25.51	25.31
R^2	0.099	0.101
Overweight (BMI ≥ 25)	0.067*	0.058*
	(0.034)	(0.030)
Mean	0.435	0.421
R^2	0.089	0.088
Obese (BMI ≥ 30)	0.035	0.035
	(0.023)	(0.023)
Mean	0.181	0.171
R^2	0.058	0.057
N	706,911	786,037

Standard errors clustered at the county level. Column 1 appends 1990-1993 BRFSS cross-sections to the original 1999–2007 BRFSS cross-sections. Column 2 appends 1990–1996 BRFSS cross-sections to the original 1999–2007 BRFSS cross-sections. Broadband variable is set to zero for all years prior to 1999 in both regressions. All regressions control for age, marital status, education (high school graduate, some college, and bachelor's degree or higher), average yearly county unemployment rate, county per capita income, county-specific linear time trends, and include month, year, and county fixed effects. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 18 Estimates of the effect of broadband availability on weight, white women samples by period

	White women: 1999–2003				White women: 2004–2007			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.808*	0.799*	0.731*	1.204**	−2.314	−2.255	−2.462	−2.501
	(0.439)	(0.429)	(0.417)	(0.573)	(2.958)	(3.008)	(2.995)	(4.706)
Mean		25.48				26.24		
R^2	0.023	0.080	0.080	0.085	0.033	0.084	0.084	0.090
Overweight (BMI ≥ 25)	0.059*	0.059*	0.058*	0.080*	0.087	0.096	0.092	0.282
	(0.034)	(0.033)	(0.033)	(0.044)	(0.164)	(0.164)	(0.165)	(0.282)
Mean		0.434				0.490		
R^2	0.020	0.072	0.072	0.077	0.028	0.077	0.077	0.082
Obese (BMI ≥ 30)	0.042	0.041	0.030	0.052	-0.053	−0.057	−0.068	0.024
	(0.026)	(0.025)	(0.025)	(0.037)	(0.156)	(0.161)	(0.161)	(0.025)
Mean		0.177				0.219		
R^2	0.017	0.046	0.046	0.052	0.026	0.055	0.055	0.060
Demographic controls		X	X	X		X	X	X
County controls			X	X			X	X
County linear time trends				X				X
N	264,457	264,457	264,457	264,457	387,170	387,170	387,170	387,170

Standard errors clustered at the county level. All regressions include month, year, and county fixed effects. Columns 2 and 6 add demographic controls: indicator variables for age, marital status, education (high school graduate, some college, and bachelor's degree or higher). Columns 3 and 7 add the county-level unemployment rate and county-level real per capita income. Columns 4 and 8 add county-specific linear time trends. All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 19 Other potential mechanisms

	White men	White women	Non-white men	Non-white women
Employed	−0.016	0.001	−0.064	0.014
	(0.030)	(0.029)	(0.109)	(0.073)
Mean	0.815	0.673	0.775	0.603
N	478,097	692,161	115,016	191,128
Currently trying to lose weight	0.150	0.019	−0.128	−0.033
	(0.124)	(0.182)	(0.107)	(0.095)
Mean	0.354	0.496	0.337	0.493
N	116,801	155.023	28,788	43,795
Has any type of health insurance	0.017	0.042*	0.020	−0.123*
	(0.026)	(0.022)	(0.095)	(0.066)
Mean	0.870	0.890	0.709	0.759
N	473,420	650,787	111,164	177,746
Routine check-up in last 12 months	0.034	−0.008	−0.207	−0.128
	(0.053)	(0.038)	(0.170)	(0.143)
Mean	0.413	0.281	0.374	0.243
N	285,671	421,964	67,626	118,157
Cholesterol check-up in last 12 months	0.132**	0.040	0.081	0.023
	(0.056)	(0.044)	(0.120)	(0.086)
Mean	0.648	0.657	0.711	0.756
N	225,196	338,167	45,664	82,122

Standard errors clustered at the county level. All regressions control for age, marital status, education (high school graduate, some college, and bachelor's degree or higher), average yearly county unemployment rate, county per capita income, county-specific linear time trends, and include month, year, and county fixed effects. Regressions for non-white samples control for race (Hispanic, other race/ethnicities; non-Hispanic black as base group). All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 20 Effect of broadband coverage on negative health behavior index

	White men	White women	Non-white men	Non-white women
Broadband_{ct}	0.021	0.078	0.183	0.212**
	(0.088)	(0.106)	(0.146)	(0.084)
N	446,694	652,851	106,927	180,696

We created an index of health behaviors such that a positive coefficient reflects more drinking and smoking and less exercise. Standard errors clustered at the county level. All regressions control for age, marital status, education (high school graduate, some college, and bachelor's degree or higher), average yearly county unemployment rate, county per capita income, county-specific linear time trends, and include month, year, and county fixed effects. Regressions for non-white samples control for race (Hispanic, other race\ethnicities; black as base group). All regressions are weighted using the BRFSS sampling weights

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

References

Akatsu H, Kuffner J (1998) Medicine and the Internet. West J Med 169(5):311–317
Akee R, Spilde K, Taylor J (2015) The Indian gaming regulatory act and its effect on American Indian economic development. J Econ Perspect 29(3):185–208
Akerman A, Gaarder I, Mogstad M (2015) The skill complementarity of broadband Internet. Q J Econ 130(4): 1781–1824
Almond D, Hoynes H, Schanzenbach D (2011) Inside the war on poverty: the impact of food stamps on birth outcomes. Rev Econ Stat 93(2):387–403
Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL (2015) Access to care and use of the Internet to search for health information: results from the US National Health Interview Survey. J Med Internet Res 17(4):e106
Anderson ML (2008) Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry preschool, and early training projects. J Am Stat Assoc 103(484):1481–1495
Atasoy H (2013) The effects of broadband Internet expansion on labor market outcomes. Ind Labor Relat Rev 66:315–345
Bailey MJ, Collins WJ (2011) Did improvements in household technology cause the baby boom? Evidence from electrification, appliance diffusion and the Amish. Am Econ J: Macroecon 3:189–217
Bailey MJ, Goodman-Bacon A (2015) The war on poverty's experiment in public medicine: community health centers and the mortality of older Americans. Am Econ Rev 105(3):1067–1104
Baker LB, Wagner TH, Singer S, Bundorf MK (2003) Use of the Internet and e-mail for health care information: results from a national survey. J Am Med Assoc 289(18):2400–2406
Bellou A (2015) The impact of Internet diffusion on marriage rates: evidence from the broadband market. J Popul Econ 28(2):265–297
Berland G, Elliot M, Morales L, Algazy J, Kravitz R, Broder M, Kanouse D, Muñoz J, Puyol J, Lara M, Watkins K, Yang H, McGlynn E (2001) Health information on the Internet: accessibility, quality, and readability in English and Spanish. J Am Med Assoc 285(20):2612–2621
Bessière K, Pressman S, Kiesler S, Kraut R (2010) Effects of Internet use on health and depression: a longitudinal study. J Med Internet Res 12(1):e6
Bhattacharya J, Bundorf MK (2009) The incidence of the healthcare costs of obesity. J Health Econ 28:649–658
Bhuller M, Havnes T, Leuven E, Mogstad M (2013) Broadband Internet: an information superhighway to sex crime? Rev Econ Stud 80:1237–1266
Biermann J, Golladay G, Greenfield M, Baker L (1999) Evaluation of cancer information on the Internet. *Cancer* 86:381–390
Carrell SE, Hoekstra M, West JE (2011) Is poor fitness contagious? Evidence from randomly assigned friends. J Public Econ 95(7-8):657–663
Case A, Paxson C (2008) Stature and status: height, health and labor market outcomes. J Polit Econ 116(3): 499–532
Cawley JH (2011) The economics of obesity. In: Cawley JH (ed) The Oxford handbook of the social science of obesity. Oxford University Press, New York, pp 303–312
Cawley J (2015) An economy of scales: a selective review of obesity's economic causes, consequences, and solutions. J Health Econ 43:244–268
Cawley J, Meyerhoefer C (2012) The medical care costs of obesity: an instrumental variables approach. J Health Econ 31:219–230
Cawley J, Ruhm CJ (2011) The economics of risky health behaviors. Handbook of health economics. Vol. 2. Elsevier, 95-199.
Christakis N, Fowler J (2007) The spread of obesity in a large social network over 32 years. N Engl J Med 357:370–379
Christensen GS, Miguel E (2016) Transparency, reproducibility, and the credibility of economics research. National Bureau of Economic Research Working Paper No. 22989
Courtemanche C (2009) Rising cigarette prices and rising obesity: coincidence or unintended consequence? J Health Econ 28(4):781–798
Culver JD, Gerr R, Frumkin H (1997) Medical information on the Internet: a study of an electronic bulletin board. J Gen Intern Med 12(8):466–470
Cunningham SA, Vaquera E, Maturo CC, Venkat Narayan KM (2012) Is there evidence that friends influence body weight? A systematic review of empirical research. Soc Sci Med 75:1175–1183

DellaVigna S, Ethan K (2007) The Fox News effect: media bias and voting. Q J Econ 122(3):1187–1234
Dettling LJ (2017) Broadband in the labor market: the impact of residential high-speed Internet on married women's labor force participation. Ind Labor Relat Rev 70(2):451–482
Dhurandhar EJ, Kaiser KA, Dawson JA, Alcorn AS, Keating KD, Allison DB (2015) Predicting adult weight change in the real world: a systematic review and meta-analysis accounting for compensatory changes in energy intake or expenditure. Int J Obes. 39(8):1181–1187
Donald A, Lindenberg B, Humphreys L (1998) Medicine and health on the Internet: the good, the bad, and the ugly. J Am Med Assoc 280(15):1303–1306
Finkelstein EA, Trogdon JG, Cohen JW, Dietz W (2009) Annual medical spending attributable to obesity: payer- and service-specific estimates. Health Aff 28(5):w822–w831
Fletcher J (2011) Peer effects and obesity. In: Cawley JH (ed) The Oxford handbook of the social science of obesity. Oxford University Press, New York, pp 303–312
Gentzkow M (2006) Television and voter turnout. Q J Econ 121(3):931–972
Gentzkow M, Shapiro JM (2008) Preschool television viewing and adolescent Test Scores: Historical evidence from the Coleman study. Q J Econ 123(1):279–323
Guldi M, Herbst C (2017) Offline effects of online connecting: the impact of broadband diffusion on teen fertility decisions. J Popul Econ 30(1):69–91
Hoynes H, Schanzenbach D (2009) Consumption responses to in-kind transfers: evidence from the introduction of the food stamp program. Am Econ J: Appl Econ 1(4):109–139
Hoynes H, Page M, Stevens AH (2011) Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program. J Public Econ 95(7–8):813–827
Impicciatore P, Pandolfini C, Casella N, Bonati M (1997) Reliability of health information for the public on the world wide web: systematic survey of advice on managing fever in children at home. Br Med J 314: 1875–1879
Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel DH, Rickman AD, Wahed AS, Belle SH (2016) Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. J Am Med Assoc **316**(11):1161–1171
Jensen R, Oster E (2009) The power of TV: cable television and women's status in rural India. Q J Econ 124(3):1057–1094
Jerome GJ, Myers VH, Young DR, Matthews-Ewald MR, Coughlin JW, Wingo BC, Ard JD, Champagne CM, Funk KL, Stevens VJ, Brantley PJ (2015) Psychosocial predictors of weight loss by race and sex. Clin Obes 5(6):342–348
Kearney M, Levine P (2015a) Media influences on social outcomes: the impact of MTV's 16 and pregnant on teen childbearing. Am Econ Rev 105(12):3597–3632
Kearney M, Levine P (2015b) Early childhood education by MOOC: Lessons from Sesame Street. National Bureau of Economic Research Working Paper No. 21229
Kling JR, Liebman JB, Katz LF (2007) Experimental analysis of neighborhood effects. Econometrica 75(1): 83–119
Kolko J (2012) Broadband and local growth. J Urban Econ 71:100–113
Kremer M, Levy D (2008) Peer effects and alcohol use among college students. J Econ Perspect **22**(3):189–206
La Ferrara E, Chong A, Duryea S (2012) Soap operas and fertility: evidence from Brazil. Am Econ J: Appl Econ 4(4):1–31
Lakdawalla D, Philipson T, Bhattacharya J (2005) Welfare-enhancing technological change and the growth of obesity. Am Econ Rev: Pap Proc 179th Ann Meet Am Econ Assoc 95(2):253–257
List, JA, Shaikj AM, Xu Y (2016) Multiple hypothesis testing in experimental economics. National Bureau of Economic Research Working Paper No. 21875
Lundborg P (2006) Having the wrong friends? Peer effects in adolescent substance use. J Health Econ 25:214–233
McLellan F (1998) Like hunger, like thirst': patients, journals, and the Internet. Lancet 352:S39–S43
Melanson EL, Keadle SK, Donnelly JE, Braun B, King NA (2013) Resistance to exercise-induced weight loss: compensatory behavioral adaptations. Med Sci Sports Exerc 45:1600–1609
Miller WC, Koceja DM, Hamilton EJ (1997) A meta-analysis of the past 25 years of weight loss research using diet, exercise or diet plus exercise intervention. Int J Obes Relat Metabol Disord 21:941–947
Owen N, Healy GN, Matthews CE, Dunstan DW (2010) Too much sitting: the population health science of sedentary behavior. Exerc Sport Sci Rev 38(3):105–113
Potarca G (2017) Does the Internet affect assortative mating? Evidence from the U.S. and Germany. Soc Sci Res 61:278–297

Poy S, Schüller S (2016) Internet and voting in the Web 2.0 era: evidence from a local broadband policy. *IZA* Discussion Paper No. 9991
Purcell G, Wilson P, Delamothe T (2002) The quality of health information on the Internet: as for any other medium it varies widely; regulation is not the answer. Br Med J **324**:557–558
Ross SL et al (2008) Mortgage lending in Chicago and Los Angeles: a paired testing study of the pre-application process. J Urban Econ 63(3):902–919
Ruhm C (2005) Health living in hard times. J Health Econ **24**(2):341–363
Stigler G (1961) The economics of information. J Polit Econ 69(3):213–225
Suiziedelyte A (2012) How does searching for health information on the Internet affect individuals' demand for health care services? Soc Sci Med 75(10):1828–1835
Thomas DM, Bouchard C, Church T, Slentz C, Kraus WE, Redman LM, Martin CK, Silva AM, Vossen M, Westerterp K, Heymsfield SB (2012) Why do individuals not lose more weight from an exercise intervention at a defined dose? An energy balance analysis. Obes Rev 13(10):835–847
Trudeau J (2015) the role of new media on teen sexual behaviors and fertility outcomes—the case of '16 and Pregnant. South Econ J 82(3):975–1003
Wagner TH, Hu T, Hibbard JH (2001) The demand for consumer health information. J Health Econ 20:1059–1075
Wellman B, Gulia M (1999) Net surfers don't ride alone: virtual community as community. In: Wellman B (ed) Networks in the global village. West-view, Boulder, pp 331–367
Wellman B, Haase AQ, Witte J, Hampton K (2001) Does the Internet increase, decrease or supplement social capital? Am Behav Sci 45(3):436–455
Wellman, B, Salaff, J, Dimitrova, D, Garton, L, Gulia, M, Haythornthwaite, C. (1996) Computer networks as social networks: collaborative work, telework, and virtual community. Annual Review of Sociology 22: 213-238
Zhao S (2006) Do Internet users have more social ties? A call for differentiated analyses of Internet use. J Comput Mediat Commun 11(3):844–862