Down the Rabbit Hole: Habit-formation in Internet Use among Unemployed Workers*

Tristan Potter†
Drexel University

April 4, 2022

Abstract

This paper tests for habit-formation in leisure-related internet use (LIU) using time-diary data from a panel of unemployed workers. Drawing on insights from the consumption-habit literature, I use a model of intertemporal time allocation to derive a test for habit-formation in leisure activities. The data reveal strong evidence of habit-formation in LIU among the Generation-X age cohort. With the exception of reading, I find no evidence of habit-formation in offline leisure.

Keywords: Habit-formation; leisure; internet use
JEL Classification: D91, J22

*Thanks to the editor and an anonymous referee for helpful comments and suggestions. Thanks also to Sebastien Bradley, Kamyar Kamyar, Ioannis Kospentaris, Andre Kurmann, Chris Laincz, Irina Murtazashvili, Maria Olivero, Ricardo Serrano-Padial, Mark Stehr, and Yoto Yotov.

†School of Economics, LeBow College of Business, Drexel University. Email: tristan.l.potter@drexel.edu. Phone: 215-895-2540. Mailing address: 3220 Market St., Philadelphia, PA 19104.

1 Introduction

There is growing concern about the large amount of free time we spend online: Congress recently considered legislation to reduce social media use;[1] the WHO now lists digital gaming as an addictive disorder;[2] and "Rabbit Hole," the *New York Times'* podcast, examines how immersive YouTube viewing can cause political radicalization.[3] These examples lead to a fundamental question about internet use: Is online leisure habit-forming? Evidence on this question is limited and typically based on neurological studies or cross-sectional analysis.[4]

Unemployment represents a unique opportunity to study habit-formation in time use. This paper uses weekly time-diary data from a panel of unemployed workers to provide causal evidence on habit-formation in leisure-related internet use (LIU)—reflecting activities such as using social media, online gaming, and video streaming. I find strong evidence of habit-formation in LIU among the Generation-X age cohort, but no such evidence among other age cohorts nor among individuals engaging in most *offline* leisure activities. The magnitude of the effects I document, and the implications for time allocation during unemployment, are substantial: A shock that initially increases LIU by one hour per day will eventually result in a member of the Generation-X cohort spending an additional 36 minutes per day engaged in LIU relative to those for whom LIU is not habit-forming.

The paper is organized as follows: Section 2 presents the model and estimating equation, Section 3 discusses data and identification, Section 4 presents results, and Section 5 concludes.

2 Model

In this section, I use a simple model of intertemporal time allocation to derive an estimating equation to test for habit-formation in leisure activities. Noting the similarities between intertemporal leisure and consumption choice, I adapt Deaton's (1992) model of habit-formation in consumption to leisure.

Consider an unemployed worker who allocates time to either some leisure activity (e.g., LIU) or stochastic time commitments (e.g., chores) that can be transferred intertemporally at cost ζ but must eventually be completed.[5] In each period t, worker i chooses how much time to devote to the leisure activity, $l_{i,t}$, to maximize

$$E_t\left[\sum_{d=0}^{D}\beta^d u(\tilde{l}_{i,t+d};\xi_{i,t+d})\right], \tag{1}$$

where β is the discount factor, $\xi_{i,t+d}$ is a preference shock, and

$$\tilde{l}_{i,t+d} \equiv l_{i,t+d} - \rho S_{i,t+d-1} \tag{2}$$

is habit-adjusted leisure where ρ governs the extent of habit-formation and $S_{i,t+d-1}$ is the habit stock,

[1] The 2019 Social Media Addiction Reduction Technology Act proposed a daily 30-minute limit on social media apps.
[2] https://www.who.int/news-room/q-a-detail/addictive-behaviours-gaming-disorder.
[3] https://www.nytimes.com/column/rabbit-hole.
[4] See Allcott et al. (2021) for a recent exception.
[5] We both "borrow" leisure (delaying chores until the weekend) and "save" it (preparing meals on Sunday).

which decays at rate δ. The habit stock evolves according to

$$S_{i,t+d} = (1-\delta)S_{i,t+d-1} + l_{i,t+d}. \tag{3}$$

The utility function u in (1) is assumed to be increasing and strictly concave in $\tilde{l}_{i,t+d}$, implying that the marginal utility of leisure is increasing in the habit stock $S_{i,t+d-1}$ when $\rho > 0$—the hallmark feature of habit-formation. Optimization is subject to three constraints:

$$l_{i,t} \leq T \tag{4}$$

$$l_{i,t} + \tau_{i,t} \leq T + b_{i,t} - (1+\zeta)b_{i,t-1} \tag{5}$$

$$b_{i,D} \leq 0, \tag{6}$$

where T denotes total time available in a period, $\tau_{i,t} \geq 0$ denotes a period-t time commitment shock, and $b_{i,t}$ denotes "borrowed" time in period t. The constraint in (4) indicates that leisure cannot exceed total time in a day; (5) indicates that leisure can be borrowed or saved to smooth marginal utility (as with consumption); and (6) indicates that time commitments must eventually be completed.

In Appendix A, I show that the optimal choice of $l_{i,t}$ implies[6]

$$\Delta l_{i,t} = \gamma_0 + \tilde{\rho}\Delta l_{i,t-1} + \nu_{i,t}, \tag{7}$$

where Δ is the first-difference operator,[7] γ_0 is a constant that depends on β and ζ, $\nu_{i,t}$ is the reduced-form error (which I discuss in Section 3.3), and $\tilde{\rho}$ measures the strength of habit-formation. If there is habit-formation (i.e. $\rho > 0$), then $\tilde{\rho} = \rho + 1 - \delta > 0$; if there is not (i.e. $\rho = 0$), then $\tilde{\rho} = 0$.[8] Thus, a finding of $\tilde{\rho} > 0$ implies habit-formation.

Intuitively, (7) reflects that, if LIU is habit-forming, then a shock that increases LIU today will increase the future marginal utility from LIU. This leads to further increases in LIU in the future, as individuals seek to smooth *habit-adjusted* LIU (rather than LIU itself). Figure 1 depicts this process.

3 Data and Identification

3.1 Data

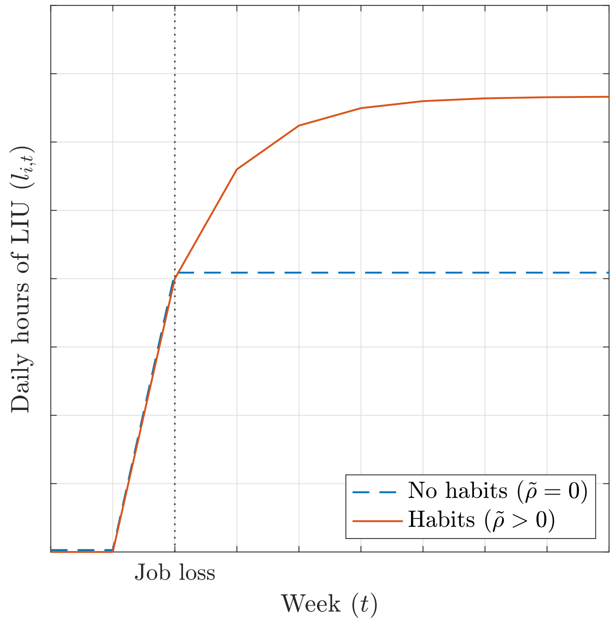
I estimate (7) on weekly panel data from the Survey of Unemployed Workers in New Jersey (SUWNJ). The SUWNJ is a once-weekly survey, lasting up to 24 weeks, of 6,025 unemployment insurance benefit recipients in New Jersey, that was conducted between the fall of 2009 and early 2010. The survey contains a detailed time-diary component in which respondents provide an hour-by-hour account of the preceding day, selecting up to two activities from 21 choices for each of 16 hours (7am-10pm).

[6] The derivation assumes an exponential (CARA) form for u, as is often used in the literature (e.g. Hayashi (1985)). A linear-quadratic form yields an analogous equation.

[7] I follow the consumption-habit literature (e.g., Deaton (1992)) to derive (7) from workers' intertemporal optimality condition. The closely-related rational addiction literature usually invokes perfect certainty to derive an estimating equation from the *intra*temporal optimality condition, resulting in an equation in levels.

[8] In the case of habit-formation, where $\tilde{\rho} = \rho + 1 - \delta$, ρ governs the effect of the previous week's leisure and $1-\delta$ the effect of leisure in all earlier weeks. However, ρ and δ are not separately identified without more assumptions. See Appendix A.

Figure 1: Habit-formation



Following Krueger and Mueller (2011), I exclude irregular time-diary entries with either (i) no activity reported for three or more hours of a day or (ii) fewer than four distinct activities reported in a day. See Appendix B for further details.

3.2 Measuring LIU

Measuring LIU requires information on internet use and leisure. The SUWNJ activity category "Computer/Internet/Email" (CIE) provides a measure of internet use, but some CIE may not be leisure-related.[9] This observation leads to two ambiguities: (i) If a respondent *only* reports CIE during a particular hour, there is no second activity to help determine whether the time was leisure, and (ii) if a respondent reports CIE and another activity, the two activities could have occurred sequentially (e.g., sleeping followed by reading news online) or concurrently (e.g., job search using the internet)—an important distinction, since LIU is strictly positive in the former case but presumably zero in the latter.

To address these two ambiguities, I make several assumptions that guide my definition of LIU:

A1. If a respondent reports only CIE and no other activity, the full hour was LIU.[10]

A2. If a respondent reports CIE and a second activity that *typically occurs online*, then the activities occurred *concurrently* and

   (i) If the activity was leisure-related, the full hour was LIU.

[9]Data from Global Web Index (GWI) suggest that between 2012 and 2017, 78% of internet use was leisure-related (https://blog.gwi.com/chart-of-the-day/social-media-captures-30-of-online-time/).

[10]This assumption is motivated by the time-use statistics in footnote 9 and the fact that the SUWNJ has very broad work-related activity categories (e.g., "Working") but no equivalent "Leisure" category. Thus, a respondent using the internet for broadly-defined work will likely report "CIE" and "Working," while a respondent using it for leisure is likely to only report "CIE."

(ii) If the activity was non-leisure-related, none of the hour was LIU.

A3. If CIE occurs in the same hour as an activity that is *not typically online*, then the activities occurred *sequentially* and approximately half of the hour was LIU.

Formally, for any hour of the day $\tau \in \{\text{7am, 8am,..., 10pm}\}$, let $a_{\tau,1}$ and $a_{\tau,2}$ denote the two reported activities.[11] Without loss of generality, order activities such that, if selected, CIE is activity $a_{\tau,1}$. Then, denoting time spent on LIU as $l_{i,t}^{\text{LIU}}$,

$$l_{i,t}^{\text{LIU}} \equiv \sum_{\tau=\text{7am}}^{\text{10pm}} \begin{cases} 60 & \text{if } a_{\tau,1} = \text{CIE} \quad \& \quad a_{\tau,2} \in \{\mathbb{L}^{\text{Online}}, \text{CIE}, \cdot\} \\ 30 & \text{if } a_{\tau,1} = \text{CIE} \quad \& \quad a_{\tau,2} \notin \{\mathbb{NL}^{\text{Online}}\} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where I define the set of leisure and non-leisure activities that typically occur online as

$$\begin{aligned} \mathbb{L}^{\text{Online}} &= \{\text{TV, Phone}\} \\ \mathbb{NL}^{\text{Online}} &= \{\text{Search, Work, Training, Course}\} \end{aligned}$$

and '·' indicates non-response. See Appendix B for further discussion.

3.3 Empirical strategy

Equation (7) holds regardless of the rate of habit decay, δ. However, δ will affect the error in (7): With slow decay, time-t marginal utility will depend on leisure both in $t-1$ and in earlier periods. I discuss the implications for identification below.

Consider first the case of quickly decaying habits ($\delta \to 1$). In this case, the error in (7) is

$$\nu_{i,t} = \xi_{i,t} - \xi_{i,t-1} + e_{i,t}, \tag{9}$$

where $e_{i,t}$ is forecast error. Because $l_{i,t-1}$ is correlated with $\xi_{i,t-1}$, $Cov(\Delta l_{i,t-1}, \nu_{i,t}) \neq 0$, so OLS will produce inconsistent estimates.

Accordingly, I first follow Anderson and Hsiao (1981) and instrument $\Delta l_{i,t-1}$ with $l_{i,t-2}$ ("A-H"). Because $l_{i,t-2}$ is highly correlated with $\Delta l_{i,t-1}$ but uncorrelated with $\Delta \xi_{i,t}$, this will alleviate the problem described above. If $l_{i,t}$ is measured with error, however, A-H estimates will be attenuated towards zero.

To address this attenuation, I propose a second instrument: twice-lagged *total* leisure time ("TLT"). This instrument is correlated with $\Delta l_{i,t-1}$ by construction, because LIU is a component of leisure. Furthermore, if coarser divisions of time are easier for respondents to recall than more granular ones—for example, if it is easier for a respondent to remember that an online course ended at 8pm, after which they enjoyed leisure until going to bed, than it is to remember exactly how that leisure time was divided—then aggregating various leisure activities will reduce measurement error. Thus, the proposed instrument accepts a higher variance (due to a weaker correlation with $\Delta l_{i,t-1}$) in exchange for reducing bias, so we should expect higher but less precise point estimates.

[11] Subscripts 1 and 2 are arbitrary and do not reflect the timing of activities.

If habits decay slowly ($\delta < 1$) and additional covariates do not fully capture preference shocks, then the error in (7) becomes

$$\nu_{i,t} = \xi_{i,t} - (2-\delta)\xi_{i,t-1} + (1-\delta)\xi_{i,t-2} + e_{i,t} - (1-\delta)e_{i,t-1}. \tag{10}$$

In Appendix A, I show that the presence of $(1-\delta)\xi_{i,t-2}$ in $\nu_{i,t}$ will attenuate point estimates from both instruments. Thus, positive point estimates represent particularly compelling evidence of habit-formation.

4 Results

4.1 Leisure-related internet use

Table 1 reports results from estimation of (7) for three generational cohorts: Millennials (20-29), Generation-X (30-49) and Baby Boomers (50-69). All regressions include controls for jobless duration, weekend interviews, sex, education, savings, marital status, and children, as well as calendar-week and survey-cohort dummies.[12]

Table 1: Habit-formation in LIU

	Millennials		*Generation-X*		*Baby Boomers*	
	A-H	TLT	A-H	TLT	A-H	TLT
Habit ($\tilde{\rho}$)	-0.11	0.02	0.13***	0.37***	0.06*	0.08
	(0.14)	(0.20)	(0.05)	(0.14)	(0.04)	(0.06)
Full controls:	×	×	×	×	×	×
A-B test: AR(2)†	0.95	0.61	0.85	0.24	0.08	0.11
A-B test: AR(3)†	0.65	0.59	0.85	0.80	0.33	0.33
K-P test‡	0.00	0.00	0.00	0.00	0.00	0.00
Observations	784	784	4,897	4,897	7,880	7,880

†p-values; H_0: No order-i autocorrelation in residuals.
‡p-values; H_0: Model underidentified.
Robust standard errors in parentheses.
Notes: Results include month and survey-cohort dummies and controls for sex, education, unemployment duration, savings, marital status and children.

The results indicate strong evidence of habit-formation in LIU for Generation-X. Interestingly, however, we see no such evidence for Millennials and only limited evidence for Baby Boomers—I return to this observation below. Concretely, $\tilde{\rho} = 0.37$ implies that a shock that initially increases LIU by one hour will eventually result in individuals spending *an additional* 36 minutes on LIU compared with $\tilde{\rho} = 0$.

Consistent with measurement error that is mitigated by the TLT instrument, point estimates from the TLT instrument are larger and less precise than those from the A-H instrument. Arellano and Bond's (1991) serial correlation test suggests that any attenuation due to $\delta < 1$ is limited.[13] The Kleinbergen-Paap test suggests that under-identification is not a problem with either instrument.

[12]See Appendix C for full results.
[13]See Appendix A for further discussion.

Results are robust to various alternative specifications, sample selection criteria, and definitions of LIU/TLT—see Appendix C.

What explains the discrepancy between Millennials and Generation-X? One possibility is that because Millennials are "digitally native," they are less susceptible to habit-forming features of LIU. Alternatively, habits may have become entrenched prior to unemployment for leisure-rich Millennials (who presumably have fewer family obligations, etc.). In this case, unemployment may provide less scope for identifying $\tilde{\rho}$, even if LIU is, in fact, habit-forming for these individuals.[14]

4.2 Offline leisure

Finally, I test whether the habit-formation observed in LIU is also present in other leisure activities. Table 2 reports estimates of (7) for the Generation-X sample (for whom null results are likely to represent a true lack of habits given the foregoing discussion) using various *offline* leisure activities (i.e., replacing $l_{i,t}^{\text{LIU}}$ with $l_{i,t}^{\text{TV}}$, $l_{i,t}^{\text{Reading}}$, etc.).

Table 2: Offline leisure

	TV		*Reading*		*Socializing*		*Exercise*	
	A-H	TLT	A-H	TLT	A-H	TLT	A-H	TLT
Habit ($\tilde{\rho}$)	-0.06	-0.04	0.15**	0.54*	0.04	0.08	0.05	0.41
	(0.06)	(0.09)	(0.07)	(0.29)	(0.05)	(0.27)	(0.07)	(0.35)
Full controls:	×	×	×	×	×	×	×	×
A-B test: AR(2)†	0.39	0.57	0.99	0.27	0.57	0.99	0.35	0.46
A-B test: AR(3)†	0.82	0.81	0.66	0.65	0.23	0.23	0.10	0.21
K-P test‡	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Observations	4,897	4,897	4,897	4,897	4,897	4,897	4,897	4,897

†p-values; H_0: No order-i autocorrelation in residuals.
‡p-values; H_0: Model underidentified.
Robust standard errors in parentheses.
Notes: Generation-X sample.

Table 2 reveals no evidence of habit-formation in offline leisure, with the possible exception of reading. The apparent absence of habit-formation is interesting in its own right, but the results are also important because they suggest that the results for LIU in Table 1 are not due to some mechanical feature of the data or estimation procedure. They also demonstrate that it is not screen time (such as watching TV), but rather *online* time, that is habit-forming.

5 Conclusion

This paper has used time-diary data from a panel of unemployed workers to test for habit-formation in leisure-related internet use (LIU)—activities such as social media, online gaming, and video streaming. I document robust evidence of habit-formation in LIU among unemployed workers in the Generation-X age cohort, but find no such evidence among other age cohorts nor among individuals engaging in

[14] I thank a referee for suggesting this possibility.

most offline leisure activities. The results imply large effects on the time allocation of affected workers, with potentially important implications for individuals' job search and labor supply decisions. Future research should explore these implications, as well as whether habit-formation in LIU co-exists with other behavioral mechanisms such as present bias and the implications for policies designed to affect how individuals engage with online leisure technologies.

References

Allcott, H., M. Gentzkow, and L. Song (2021). Digital addiction. *Working Paper*.

Anderson, T. and C. Hsiao (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*.

Deaton, A. (1992). *Understanding Consumption.* Oxford: Clarendon Press.

Hayashi, F. (1985). The permanent income hypothesis and consumption durability: Analysis based on japanese panel data. *Quarterly Journal of Economics*.

Krueger, A. and A. Mueller (2011). Job search, emotional well-being and job-finding in a period of mass unemployment: Evidence from high-frequency longitudinal data. *Brookings Papers on Economic Activity*.

Appendices

A Derivation of estimating equation

In this appendix, I derive the estimating equation in (7) and discuss some of the empirical implications. The model and derivation follow Deaton (1992), who studies habit-formation in consumption.

A.1 Model

Consider a generalized version of the model described in the main text in which unemployed workers derive utility from A different leisure activities, one of which is LIU.[15] In each period t, worker i chooses how much time to devote to leisure activity a, $l^a_{i,t}$, for $a = 1, .., A$, to maximize

$$E_t\left[\sum_{d=0}^{D} \beta^d u(\tilde{l}^1_{i,t+d}, \tilde{l}^2_{i,t+d}, ..., \tilde{l}^A_{i,t+d}; \xi^1_{i,t+d}, \xi^2_{i,t+d}, ..., \xi^A_{i,t+d})\right] \tag{A.1}$$

where β is the discount factor, $\xi^a_{i,t+d}$ is a taste shock to leisure activity a, and

$$\tilde{l}^a_{i,t+d} \equiv l^a_{i,t+d} - \rho^a S^a_{i,t+d-1} \tag{A.2}$$

is habit-adjusted leisure. $S^a_{i,t-1}$ is the leisure habit stock, defined as

$$S^a_{i,t-1} = (1-\delta^a)S^a_{i,t-2} + l^a_{i,t-1}, \tag{A.3}$$

and δ^a governs the rate of decay of the habit stock and ρ^a governs the extent of habit-formation for activity a. Notice that this model of habit-formation (besides being related to leisure rather than consumption) is a special case of the model of habit-formation described in Section 1.3 of Deaton (1992). Specifically, Deaton's model nests the current model when there is a single leisure activity (and hence we do not need a superscripts) via $\theta = \delta$, $\alpha = 1 - \beta$, and $\beta = -\rho/(1-\delta)$. Optimization is subject to three constraints:

$$\textstyle\sum_{a=1}^{A} l^a_{i,t} \leq T \tag{A.4}$$

$$\textstyle\sum_{a=1}^{A} l^a_{i,t} + \tau_{i,t} \leq T + b_{i,t} - (1+\zeta)b_{i,t-1} \tag{A.5}$$

$$b_{i,D} \leq 0 \tag{A.6}$$

where T denotes the total time available in any period, $\tau_{i,t} \geq 0$ denotes an idiosyncratic time commitment shock in period t that detracts from total (lifetime) time available for leisure, $b_{i,t}$ denotes "borrowed" time in period t, and ζ is the intertemporal price of leisure, analogous to the interest rate for consumption. The constraint in (A.4) indicates that leisure cannot exceed total time in a day, (A.5) indicates that leisure can be borrowed (or saved) in a manner largely analogous to consumption—reflecting the intuition described in Section 2[16]—and (A.6) indicates that all time

[15]The model described below generalizes the model in the main text by allowing for multiple leisure activities under the assumption that u is separable in the various activities. The assumption of separability is important and is also found in much of the consumption-habit literature.

[16]For more intuition on (A.5), let $b_{i,t-1} = 0$, rearrange, and note that the constraint will always bind, so we can write $T = \sum_{a=1}^{A} l^a_{i,t} + \tau_{i,t} - b_{i,t}$. Thus, total time in a day (T) is divided between leisure ($\sum_{a=1}^{A} l^a_{i,t}$) and non-leisure

commitments must eventually be completed.

I maintain two assumptions throughout the analysis: First, in keeping with the consumption-habit literature, I assume that the utility function is separable in the various (habit-adjusted) leisure activities.[17] Thus, without loss of generality, I proceed by focusing on the case of a single leisure activity (i.e. $A = 1$), and drop a superscripts, understanding that the analysis applies to any category of leisure. Second, I assume that the time constraint in (A.4) is slack, so that the associated Lagrange multipliers are zero. This assumption reflects the fact that in the SUWNJ virtually no respondents spend all of their time on leisure.[18]

The recursive nature of the habit stock $S_{i,t}$ means that it will be useful to write the optimization problem recursively. To do this, solve (A.5) for $l_{i,t}$ (noting that (A.5) will always bind and that we are focusing on a single leisure activity and have thus dropped a superscripts), substitute the resulting equation into (A.2), and substitute that resulting equation into (A.1). Furthermore, temporarily drop i subscripts to ease notation. Letting $V_t(b_{t-1}, S_{t-1})$ denote the value of entering period t with "debt" b_{t-1} and habit stock S_{t-1}, we can write (A.1) recursively as

$$V_t\Big(b_{t-1}, S_{t-1}\Big) = \max_{b_t} \Bigg\{ u\Big(T - \tau_t + b_t - (1+\zeta)b_{t-1} - \rho S_{t-1}\Big) + \beta E_t V_{t+1}\Big(b_t, (1-\delta)S_{t-1} + T - \tau_t + b_t - (1+\zeta)b_{t-1}\Big) \Bigg\}. \tag{A.7}$$

The first-order condition for b_t is

$$u'_t = -\beta E_t V_{1,t+1} - \beta E_t V_{2,t+1}, \tag{A.8}$$

where u'_t denotes the derivative of the period-t utility function and $V_{k,t+1}$ denotes the derivative of the value function V_{t+1} with respect to its k^{th} argument. The two envelope conditions are

$$V_{1,t} = -(1+\zeta)u'_t - \beta(1+\zeta)E_t V_{2,t+1} \tag{A.9}$$

$$V_{2,t} = -\rho u'_t + (1-\delta)\beta E_t V_{2,t+1}. \tag{A.10}$$

Combining (A.8), (A.9) and (A.10), we can eliminate references to the value function to obtain

$$u'_t = \beta E_t u'_{t+1}(\rho + 1 - \delta + 1 + \zeta) - (1+\zeta)(\rho + 1 - \delta)\beta^2 E_t u'_{t+2}. \tag{A.11}$$

Continuing to follow Deaton (1992), using the forward-shift operator F, defined by $Fx_t = x_{t+1}$, we can rearrange the preceding and factor as

$$u'_t\left[1 - \beta(1+\zeta)E_t F\right]\left[1 - (\rho + 1 - \delta)\beta E_t F\right] = 0. \tag{A.12}$$

activities ($\tau_{i,t} - b_{i,t}$), where the latter is just total new time commitments (e.g., needing to clean the house) net of borrowed time (e.g., cleaning put off until the future).

[17] The consumption-habit literature typically assumes separability of the utility function in various types of consumption goods (and implicitly leisure) in order to estimate models using data on, e.g., food expenditures.

[18] 99% of time diaries show fewer than 10 hours of leisure, out of a total possible 16 hours, which suggests that the vast majority of the sample devotes a significant amount of time each day to non-leisure activities. This is consistent with the constraint in (A.4) not binding.

The solution associated with the second root can be ruled out as T becomes large,[19] leaving only the solution associated with the first root,

$$u'_t = \beta(1+\zeta)E_t u'_{t+1} \tag{A.13}$$

which is precisely the condition derived in Deaton (1992). Returning to the use of i subscripts, we can rewrite (A.13) as

$$\beta(1+\zeta)\frac{u'_{i,t}}{u'_{i,t-1}} = 1 - \epsilon_{i,t}, \tag{A.14}$$

where $\epsilon_{i,t}$ is the forecast error which, under rational expectations, is serially uncorrelated and satisfies $E_{t-1}[\epsilon_{i,t}] = 0$. In the context of a model of time allocation such as the one considered here, these may be thought of as unexpected shocks to an individual's time budget constraint (analogous to permanent income shocks in the context of consumption), such as job loss, receiving a job offer, or having to take care of an ill family member. Taking logs, rearranging, and using the approximation $\ln(1-\epsilon_{i,t}) \approx -\epsilon_{i,t}$,[20] this is

$$\ln(u'_{i,t-1}) - \ln(u'_{i,t}) = \ln(\beta(1+\zeta)) + \epsilon_{i,t}. \tag{A.15}$$

To facilitate deriving an estimating equation in the presence of a slowly decaying habit stock, it is convenient to assume exponential utility such that u is of the form $u(\tilde{l}_{i,t};\xi_{i,t}) = -\mu\exp\{\frac{\xi_{i,t}-\tilde{l}_{i,t}}{\mu}\}$, where $\xi_{i,t}$ is the preference shock and $\mu > 0$ is the coefficient of absolute risk aversion.[21]

This utility function implies $\ln(u'_{i,t}) = \frac{\xi_{i,t}-\tilde{l}_{i,t}}{\mu}$. Substituting this into (A.15), using (A.2), and rearranging, we arrive at

$$\Delta l_{i,t} = \mu\ln(\beta(1+\zeta)) + \rho\Delta S_{i,t-1} + \Delta\xi_{i,t} + e_{i,t}, \tag{A.16}$$

where $e_{i,t} \equiv \mu\epsilon_{i,t}$ and Δ is the first-difference operator. Notice that if there is no habit-formation (i.e., $\rho = 0$), $\Delta S_{i,t-1}$ drops out and lagged growth in leisure ($\Delta l_{i,t-1}$) will not appear on the right-hand side of (A.16), implying that if we include it in our regression we should expect a coefficient of zero. On the other hand, if $\rho > 0$, then we need to eliminate the $\Delta S_{i,t-1}$ term from (A.16). Taking the difference between (A.3) and its first lag gives

$$\Delta S_{i,t-1} = (1-\delta)\Delta S_{i,t-2} + \Delta l_{i,t-1}. \tag{A.17}$$

[19] See Hayashi (1985).

[20] Hayashi (1985) uses the same approximation in a similar derivation. As he notes, we could alternatively have imposed conditional normality to obtain the same result.

[21] Exponential (CARA) utility is common in economics and finance; Hayashi (1985) (an important reference in the consumption-habit literature) uses the same utility function that I use to derive a similar estimating equation. With other functional forms (e.g., CRRA), we will in general need to take a quadratic approximation of the utility function (or a linear approximation of marginal utility) to be able to derive an estimating equation (an exception is when habits decay fully between periods, i.e., $\delta \to 1$, in which case we can use the approximation $\ln(l_{i,t} - \rho l_{i,t-1}) \approx \ln(l_{i,t}) - \rho\ln(l_{i,t})$ to derive an estimating equation from a CRRA utility function). Thus, we could have alternatively assumed quadratic utility (as is also common in the literature), implying linear marginal utility, in which case with $\beta(1+\zeta) \approx 1$ we recover a first-order autoregression in $\Delta l_{i,t}$, as in (A.19) below.

Iterating (A.16) backward in time and solving for $\Delta S_{i,t-2}$, we have

$$\Delta S_{i,t-2} = \frac{1}{\rho}\Big[\Delta l_{i,t-1} - \mu \ln(\beta(1+\zeta)) - \Delta\xi_{i,t-1} - e_{i,t-1}\Big]. \tag{A.18}$$

Consistent with the discussion above, note that this is only well defined when $\rho \neq 0$. Finally, substituting (A.18) into (A.17) and (A.17) into (A.16), we arrive at

$$\Delta l_{i,t} = \gamma_0 + (\rho + 1 - \delta)\Delta l_{i,t-1} + \nu_{i,t} \tag{A.19}$$

where γ_0 is a constant and $\nu_{i,t}$ is the composite error term, defined, respectively, as

$$\gamma_0 \equiv \delta\mu \ln(\beta(1+\zeta)) \tag{A.20}$$
$$\nu_{i,t} \equiv \xi_{i,t} - (2-\delta)\xi_{i,t-1} + (1-\delta)\xi_{i,t-2} + e_{i,t} - (1-\delta)e_{i,t-1}. \tag{A.21}$$

Defining $\tilde{\rho} \equiv \rho + 1 - \delta$,[22] we obtain the estimating equation in the main text:[23]

$$\Delta l_{i,t} = \gamma_0 + \tilde{\rho}\Delta l_{i,t-1} + \nu_{i,t}. \tag{A.22}$$

It should be emphasized that this equation holds for any individual leisure activity, which implies that we can use it to test not only for the presence of habit-formation in LIU, but also for the presence of habit-formation in other leisure activities. Below, I elaborate on the empirical issues associated with estimation of (A.22) that I discuss in Section 3.3 of the main text.

A.2 Empirical considerations

In the main text, I estimate (A.22) using two instruments for $\Delta l_{i,t-1}$: (i) twice-lagged LIU (A-H) and (ii) twice-lagged total leisure time (TLT).[24] Both instruments are appealing, because they alleviate the inconsistency caused by the correlation between $\nu_{i,t}$ and $\Delta l_{i,t-1}$ due to the presence of $\xi_{i,t-1}$ in (A.21) and are highly correlated with $\Delta l_{i,t-1}$. Furthermore, both yield consistent estimates when $\delta = 1$ and there is no measurement error. Below, I consider the implications of these instruments when there is measurement error or $\delta < 1$.

To do so, suppose $l_{i,t}$ is measured with error in the model described above. Formally, let $l_{i,t}$ denote true leisure time and let $l^*_{i,t}$ denote measured leisure time, such that

$$l^*_{i,t} = l_{i,t} + \eta_{i,t}, \tag{A.23}$$

where $\eta_{i,t}$ is classical measurement error. Using equation (A.23), equation (A.19) can be written as

$$\Delta l^*_{i,t} - \eta_{i,t} + \eta_{i,t-1} = \gamma_0 + \tilde{\rho}\left[\Delta l^*_{i,t-1} - \eta_{i,t-1} + \eta_{i,t-2}\right] + \nu_{i,t}, \tag{A.24}$$

[22] Although ρ and δ are not separately identified from (A.19) alone, we can get traction in special cases. For example, if $\delta = 1 - \rho$ (corresponding to the case in which the direct effect of past leisure on utility decays geometrically), $\rho = \tilde{\rho}/2$.

[23] Note that a model with "durable" leisure would correspond to $\rho = -(1-\delta)$, implying $\tilde{\rho} = 0$. Thus, in this case (as well as the case discussed above in which there is simply no habit-formation, i.e., $\rho = 0$) we would expect to find no evidence of $\tilde{\rho} > 0$ when estimating (A.22).

[24] As discussed in the main text, OLS will produce inconsistent estimates of $\tilde{\rho}$ because the error term in (A.21) contains the preference shock $\xi_{i,t-1}$ which will be correlated with $\Delta l_{i,t-1}$.

which can be rearranged to obtain

$$\Delta l^*_{i,t} = \gamma_0 + \tilde{\rho}\Delta l^*_{i,t-1} + z_{i,t}, \tag{A.25}$$

where $z_{i,t}$ is the composite error term (inclusive of measurement error), defined as

$$\begin{aligned} z_{i,t} &\equiv \eta_{i,t} - (1+\tilde{\rho})\eta_{i,t-1} + \tilde{\rho}\eta_{i,t-2} + \nu_{i,t} \\ &= \eta_{i,t} - (1+\tilde{\rho})\eta_{i,t-1} + \tilde{\rho}\eta_{i,t-2} + \xi_{i,t} - (2-\delta)\xi_{i,t-1} + (1-\delta)\xi_{i,t-2} + e_{i,t} - (1-\delta)e_{i,t-1}, \end{aligned} \tag{A.26}$$

where the second equality uses the definition of $\nu_{i,t}$ in (A.21). Now, consider some instrument $m_{i,t-2}$ for $\Delta l_{i,t-2}$. Generically, the direction of the inconsistency of the 2SLS estimator of $\tilde{\rho}$, which we denote by $\hat{\tilde{\rho}}^{\text{2SLS}}$, is determined by the sign of the second term in

$$\text{plim}(\hat{\tilde{\rho}}^{\text{2SLS}}) = \tilde{\rho} + \frac{Cov(m_{i,t-2}, z_{i,t})}{Cov(m_{i,t-2}, \Delta l^*_{i,t-1})}. \tag{A.27}$$

I use this expression to study the implications of measurement error and $\delta < 1$ for the two instruments.

Measurement error

As discussed in the text, specific leisure activities are likely to be measured with error. Below, I show that measurement error will attenuate point estimates using the A-H instrument, and that the TLT instrument alleviates this problem. To focus on the implications of measurement error, assume for the moment that $\delta = 1$ (which implies that $\tilde{\rho} = \rho$).

Consider first the case of the A-H instrument: $m_{i,t-2} = l^*_{i,t-2}$. In this case, using the definition of $l^*_{i,t-2}$ in (A.23), the numerator of the second term in (A.27) becomes

$$\begin{aligned} Cov(m_{i,t-2}, z_{i,t}) &= Cov(l^*_{i,t-2}, z_{i,t}) && \text{(A.28)} \\ &= Cov(l_{i,t-2} + \eta_{i,t-2}, \eta_{i,t} - (1+\rho)\eta_{i,t-1} + \rho\eta_{i,t-2} + \nu_{i,t}) \\ &= \rho Var(\eta_{i,t-2}) > 0, && \text{(A.29)} \end{aligned}$$

where the inequality holds if there is habit-formation ($\rho > 0$). For the denominator, we have

$$\begin{aligned} Cov(l^*_{i,t-2}, \Delta l^*_{i,t-1}) &= Cov(l^*_{i,t-2}, l^*_{i,t-1}) - Var(l^*_{i,t-2}) && \text{(A.30)} \\ &\leq Var(l^*_{i,t-2}) - Var(l^*_{i,t-2}) \\ &= 0, && \text{(A.31)} \end{aligned}$$

where the second-to-last line follows from the Cauchy-Schwarz inequality. Thus, the second term on the right-hand side of (A.27) must be negative, which implies that $\text{plim}(\hat{\tilde{\rho}}^{\text{2SLS}}) < \rho$: Measurement error will attenuate A-H point estimates toward zero, thus leading to point estimates that will tend to *understate* the extent of habit-formation, as discussed in the main text.

Next, consider the TLT instrument: $m_{i,t-2} = l^*_{i,t-2} + \mathring{l}^*_{i,t-2} = (l_{i,t-2} + \eta_{i,t-2}) + (\mathring{l}_{i,t-2} + \mathring{\eta}_{i,t-2})$, where $\mathring{l}^*_{i,t-2}$ is the measured sum of *other* leisure activities, $\mathring{l}_{i,t-2}$ is the true sum of *other* leisure activities and $\mathring{\eta}_{i,t-2}$ is the corresponding measurement error. The assumption described in the main text that makes this a useful instrument—that coarser divisions of time are recorded more accurately—corresponds to an assumption that the measurement errors are negatively correlated, i.e., $Cov(\eta_{i,t-2}, \mathring{\eta}_{i,t-2}) < 0$. To see that this can eliminate the inconsistency, consider the simplest case in which $\mathring{\eta}_{i,t-2} = -\eta_{i,t-2}$. In this case, the instrument is a perfect measure of aggregate leisure time because of offsetting measurement errors, so $m_{i,t-2} = l_{i,t-2} + \mathring{l}_{i,t-2}$. Using this, the numerator of (A.27) can be written as

$$\begin{aligned} Cov(m_{i,t-2}, z_{i,t}) &= Cov(l_{i,t-2} + \mathring{l}_{i,t-2}, z_{i,t}) && \text{(A.32)} \\ &= Cov(l_{i,t-2} + \mathring{l}_{i,t-2}, \eta_{i,t} - (1+\rho)\eta_{i,t-1} + \rho\eta_{i,t-2} + e_{i,t}) \\ &= 0, && \text{(A.33)} \end{aligned}$$

where the result follows from the fact that classical measurement error is assumed to be independent of $l_{i,t-2}$ and $\mathring{l}_{i,t-2}$. Thus, under the assumptions set out in the text, the numerator of the second term in (A.27) approaches zero, so the TLT instrument eliminates the inconsistency associated with measurement error. Thus, as described in the main text, the TLT instrument should yield point estimates that are larger than the point estimates generated by the A-H instrument.

Slow decay attenuation

The second issue arises when habits decay slowly ($\delta < 1$). In this case, $(1-\delta)\xi_{i,t-2}$ now appears in the composite error term in (A.21), which may be cause for concern with instruments dated to $t-2$. However, for reasons I describe below, this issue is unlikely to be quantitatively important and, to the extent that it is, it will necessarily militate *against* a finding of habit formation. It is unlikely to be quantitatively important, because (i) by including controls in the regression, this extra term will be (partially) purged from the error term in (A.21), and (ii) the influence of $(1-\delta)\xi_{i,t-2}$ will be small when δ is large, which is consistent with the point estimates in Tables 1 and 2. Furthermore, even if habits decay slowly and the controls in the fully saturated model do not fully absorb $\xi_{i,t-2}$, its presence in the error term will *attenuate* point estimates with both instruments. Thus, the results in Tables 1 and 2 will understate the true extent of habit-formation. I elaborate on this last point below.

To ease exposition, focus on the case with no measurement error: $Var(\eta_{i,t}) = 0$. Considering first the A-H instrument ($m_{i,t-2} = l_{i,t-2}$), the numerator of the second term in (A.27) can be written as

$$\begin{aligned} Cov(m_{i,t-2}, \nu_{i,t}) &= Cov(l_{i,t-2}, \nu_{i,t}) && \text{(A.34)} \\ &= Cov(l_{i,t-2}, \xi_{i,t} - (2-\delta)\xi_{i,t-1} + (1-\delta)\xi_{i,t-2} + e_{i,t} - (1-\delta)e_{i,t-1}) \\ &= (1-\delta)Cov(l_{i,t-2}, \xi_{i,t-2}) \\ &> 0, && \text{(A.35)} \end{aligned}$$

where the inequality follows from the fact that positive shocks to the marginal utility of leisure from LIU, $\xi_{i,t-2}$, will increase time spent on LIU, $l_{i,t-2}$. The denominator of the second term in (A.27) is

$$\begin{aligned} Cov(l_{i,t-2}, \Delta l_{i,t-1}) &= Cov(l_{i,t-2}, l_{i,t-1} - l_{i,t-2}) && \text{(A.36)} \\ &= Cov(l_{i,t-2}, l_{i,t-1}) - Var(l_{i,t-2}) \\ &\leq Var(l_{i,t-2}) - Var(l_{i,t-2}) \\ &= 0, && \text{(A.37)} \end{aligned}$$

where the inequality follows again from the Cauchy-Schwarz inequality. Thus, the sign of the inconsistency is negative and the presence of slowly decaying habit will cause an underestimation of the extent of habit-formation when using the A-H instrument.

Next, consider the TLT instrument ($m_{i,t-2} = l_{i,t-2} + \mathring{l}_{i,t-2}$). To analyze this case, note first that time spent on non-LIU leisure, $\mathring{l}_{i,t-2}$, will be approximately uncorrelated with the preference shock to LIU that appears in the error term, $\xi_{i,t-2}$.[25] Thus, the numerator of the second term in (A.27) is the same as for the A-H instrument: $Cov(l_{i,t-2} + \mathring{l}_{i,t-2}, \nu_{i,t}) = Cov(l_{i,t-2}, \nu_{i,t}) > 0$ from (A.35). The sign of the denominator, on the other hand, is a priori ambiguous. Nevertheless, this is something we can directly observe in the data. The correlation between $l_{i,t-2} + \mathring{l}_{i,t-2}$ and $\Delta l_{i,t-1}$ is robustly negative, typically around -0.25. Together, a positive numerator and a negative denominator imply that, as for the A-H instrument, slow decay will tend to attenuate point estimates using the TLT instrument. Thus, even if $\delta < 1$ and the additional control variables do not eliminate $\xi_{i,t-2}$ from the error term, its presence will lead us to *underestimate* the extent of habit-formation.

Serial correlation tests

The Arellano-Bond serial correlation tests in Tables 1 and 2 provide further evidence on the issues discussed above. Inspection of equation (A.26) reveals that the presence of either measurement error ($Var(\eta_{i,t}) > 0$) or slowly decaying habits ($\delta < 1$) will impart an MA(2) structure to the error term. This, in turn, implies that the presence of either should lead to a finding of second-order autocorrelation in the residuals but no third-order autocorrelation.

The results for the TLT instrument indicate that there is not strong evidence of second- or third-order autocorrelation for Millennials or Generation-X.[26] This suggests that (i) the TLT is successfully alleviating measurement error and (ii) attenuation due to slow decay is likely not a major problem. These observations imply that the TLT instrument provides reliable estimates of habit-formation. That there is not strong evidence of second- or third-order autocorrelation with the A-H instrument also supports the conclusion that attenuation due to slow decay is not a major problem, but need not imply that there is no measurement error: Measurement error will attenuate $\hat{\hat{\rho}}^{2SLS}$ toward zero, which (via (A.26)) can lead to an erroneous failure to reject the null of no second-order autocorrelation resulting from measurement error.[27] However, the presence of attenuation from slow decay will still induce second-order autocorrelation (even with small $\hat{\hat{\rho}}^{2SLS}$), so the fact that there is not strong evidence of second-order autocorrelation for the A-H instrument provides limited evidence on measurement error, but is consistent with a limited effect of attenuation from slow decay.[28] Thus, the serial correlation tests for both instruments support the conclusion that LIU is habit-forming.

[25] If $\xi_{i,t-2}$ is anticipated, then any adjustment in $\mathring{l}_{i,t-2}$ needed to accommodate the change in $l_{i,t-2}$ in order to satisfy the lifetime time constraint (i.e. the object analogous to the lifetime budget constraint in consumption models) will occur prior to $t-2$ in order to be consistent with leisure-smoothing in non-LIU activities, implying that $Cov(\mathring{l}_{i,t-2}, \xi_{i,t-2}) = 0$. If $\xi_{i,t-2}$ is unanticipated, then for large T (i.e. as the time horizon becomes long) the required adjustment in $\mathring{l}_{i,t-2}$ to satisfy the lifetime budget constraint will tend toward zero, once again due to leisure-smoothing in non-LIU activities.

[26] The evidence against second-order autocorrelation for Baby Boomers is more tenuous, but we still cannot reject the null of no autocorrelation at conventional significance levels.

[27] This implies that it is unclear whether we should see more or less evidence of serial correlation with the A-H instrument.

[28] The difference between the A-H and TLT point estimates is consistent with measurement error that is alleviated by the TLT instrument.

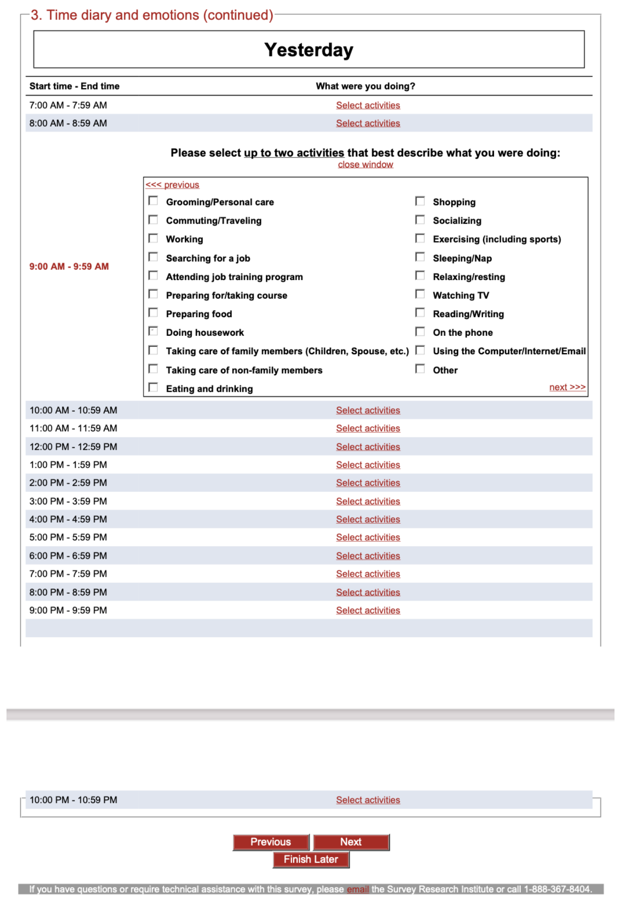
B Data

This appendix provides details on the construction of time-use variables for the empirical analysis.

B.1 Survey of Unemployed Workers in New Jersey

Each weekly survey of the SUWNJ asked respondents to complete a time diary, to account for their activities between 7am and 11pm on the previous day. Figure A1 contains a graphic of what respondents would see in this section of the survey.[29]

Figure A1: Survey of Unemployed Workers in New Jersey: Time Diary


3. Time diary and emotions (continued)

Yesterday

Start time - End time	What were you doing?
7:00 AM - 7:59 AM	Select activities
8:00 AM - 8:59 AM	Select activities
9:00 AM - 9:59 AM	Please select up to two activities that best describe what you were doing: (close window)
10:00 AM - 10:59 AM	Select activities
11:00 AM - 11:59 AM	Select activities
12:00 PM - 12:59 PM	Select activities
1:00 PM - 1:59 PM	Select activities
2:00 PM - 2:59 PM	Select activities
3:00 PM - 3:59 PM	Select activities
4:00 PM - 4:59 PM	Select activities
5:00 PM - 5:59 PM	Select activities
6:00 PM - 6:59 PM	Select activities
7:00 PM - 7:59 PM	Select activities
8:00 PM - 8:59 PM	Select activities
9:00 PM - 9:59 PM	Select activities
10:00 PM - 10:59 PM	Select activities

<<< previous

☐ Grooming/Personal care	☐ Shopping
☐ Commuting/Traveling	☐ Socializing
☐ Working	☐ Exercising (including sports)
☐ Searching for a job	☐ Sleeping/Nap
☐ Attending job training program	☐ Relaxing/resting
☐ Preparing for/taking course	☐ Watching TV
☐ Preparing food	☐ Reading/Writing
☐ Doing housework	☐ On the phone
☐ Taking care of family members (Children, Spouse, etc.)	☐ Using the Computer/Internet/Email
☐ Taking care of non-family members	☐ Other
☐ Eating and drinking	

next >>>

Previous Next
Finish Later

If you have questions or require technical assistance with this survey, please email the Survey Research Institute or call 1-888-367-8404.

Respondents thus had a total of 21 activities from which to choose. For each hour of the time diary, respondents were able to choose up to two activities to describe what they were doing. As described in the main text, I follow the procedure described in Krueger and Mueller (2011) to deal with irregular time diary entries (see their Appendix "Description of the Survey and Data Coding"). Specifically, I restrict the sample to time diaries in which at least one activity is reported in at least 14 of 16 hours of the day, and at least four distinct activities are reported throughout a given day. These restrictions help to ensure that errant time diary entries are not driving the results. Results continue to hold for the unrestricted sample—see Appendix C.

[29] Complete survey data and documentation can be obtained from: https://dss.princeton.edu/catalog/resource1350.

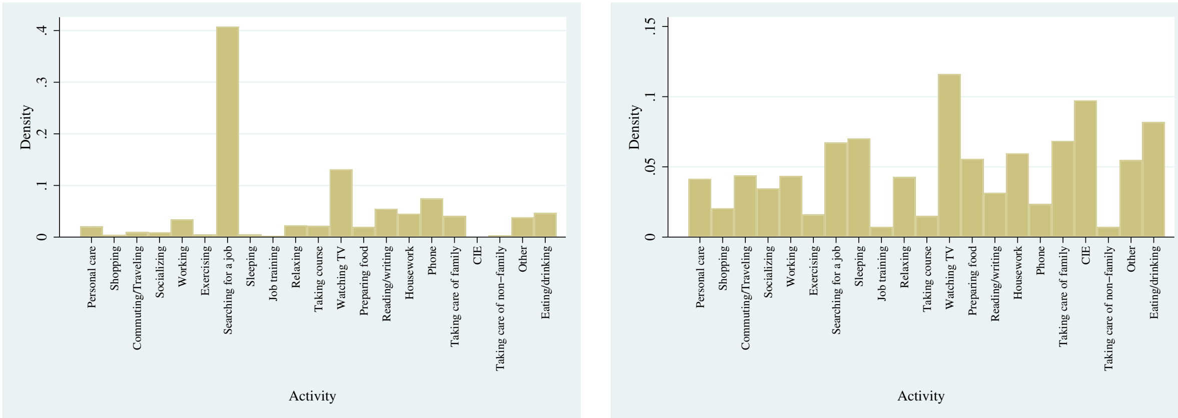
B.2 Variable construction

Leisure-related internet use (LIU)

As discussed in the main text, the definition of LIU in equation (8) is intended to resolve two ambiguities in the SUWNJ data: (i) a response of only CIE could potentially reflect some non-leisure related internet use, and (ii) if two different activities are reported during a given hour, there is no information on whether the activities were performed sequentially (e.g., "Sleeping" followed by "Eating") or concurrently (e.g., "CIE" and "Searching for a job"). I address these issues by making several assumptions (see the main text) that guide my definition of LIU in (8). Below, I briefly elaborate on each issue.

The first ambiguity results from the fact that when a respondent only reports CIE, there is no second activity to aid in determining what the internet was being used for. There are several reasons to think that such episodes primarily reflect online leisure. First, time-use data from sources other than the SUWNJ indicate that a significant majority of all time spent online is leisure-related. For example, as discussed in the main text, Global Web Index (GWI) finds that at least 78% of online time between 2012 and 2017 was leisure-related: 33% social media, 16% online TV and streaming, 16% music streaming, and 13% online press. Of course, this is not an exhaustive account of online leisure activities—for example, online gaming is notably absent—and thus 78% likely represents a lower bound on the true number. Second, the time-use activity options in the SUWNJ (see Figure A1) provide a more comprehensive account of possible work-related online time than possible leisure-related online time: "Searching for a job" and "Working" alone are sufficiently broad to cover most non-leisure related activities that might take place online, whereas it is easy to imagine common leisure activities that do not obviously fit into any of the categories in A1 (of the top four leisure categories in the GWI data above, only "online TV and streaming" seems to naturally fit into one of the SUWNJ categories). Moreover, there is no catch-all category for leisure analogous to the "Working" category. Accordingly, as described in the main text, in the baseline results I assume that if only CIE is reported the full hour is LIU. To ensure that this assumption is not driving the results, in Appendix C I report results for several alternative definitions of LIU, one of which allocates 45 minutes to LIU if a respondent only reports CIE (corresponding to roughly 78% of a full hour, as in the GWI data) rather than a full hour.

The second ambiguity stems from the fact that even if a respondent reports a second activity along with CIE, there is no information in the SUWNJ to determine whether the activities occurred concurrently or sequentially. This distinction matters, because hours in which leisure activities occur concurrently with CIE (e.g., streaming TV) should be recorded as a full hour of LIU and hours in which non-leisure activities occur concurrently with CIE (e.g., online job search) should be recorded as no LIU. On the other hand, hours in which CIE occurs sequentially with other activities should presumably be recorded as 30 minutes of LIU (per the discussion in the preceding paragraph and assuming the other activities occur offline). A simple way to assess whether respondents who report two activities are, in some cases, reporting two activities that occurred concurrently with CIE (rather than sequentially) is to look at the distribution of activities that occur conditional on also engaging in CIE in the same hour. If respondents only report activities sequentially, we should not expect to see a systematic relationship between activities that frequently occur online (such as watching TV or job search) and activities that occur in the same hour as CIE. Figure A2 reports both the



(a) Activities in same hour as CIE

(b) All activities

Figure A2: Distribution of activities

distribution of activities that occur in the same hour as CIE and the unconditional distribution of activities. Inspection of the two panels suggests a substantial degree of reporting of concurrent activities: The activities that occur most frequently in the same hour as CIE are "Searching for a job," "Watching TV," and "Phone," all of which frequently occur *while* online. This observation motivates two components of the definition of LIU in (8): Allocating a full hour to LIU when a respondent reports engaging in CIE and also "Watching TV" or "Phone," but also allocating no time to LIU when a respondent reports engaging in CIE and also "Searching for a job" (and several other activities that frequently occur online).

Offline leisure activities

The offline leisure activities considered in Table 2 correspond, respectively, to "Watching TV," "Reading/Writing," "Socializing," and "Exercising (including sports)" in Figure A1. The variables are constructed in a manner broadly similar to LIU, although construction is somewhat more straightforward because issues associated with identifying online leisure time are not as prominent. Specifically, I allocate 60 minutes to an activity if a respondent reports only engaging in that activity in a particular hour, 30 minutes to an activity if a respondent reports engaging in that activity and a second (different) activity in a particular hour,[30] and zero minutes to an activity otherwise. Given the goal of focusing on *offline* leisure, for the first three activities ("Watching TV," "Reading/Writing," and "Socializing"), I allocate zero minutes to the activity if it occurs in the same hour as CIE, which reflects a concern that these activities may have happened concurrently with internet use (e.g., reading news online or socializing in a chatroom) rather than sequentially.[31] This is clearly imperfect in light of the discussion of concurrent versus sequential activities above—for instance, if a respondent reads a book offline for half of an hour and then spends the remainder of the hour browsing the internet, I will erroneously assign zero minutes to "Reading/Writing." Reassuringly, however, this assumption regarding CIE is not important for the results: If I instead assume that no activities occur concurrently with CIE and thus simply allocate 30 minutes to each activity if it is reported in

[30] An exception to this is when CIE is the second activity; see below.

[31] On the other hand, "Exercise" is unlikely to happen at the same time as CIE.

the same hour as another activity (including CIE), the results are essentially unchanged.

Total leisure time (TLT)

Defining total leisure time (TLT) is complicated by ambiguity concerning which of the activity categories included in the SUWNJ are properly understood as leisure. To address this ambiguity in a way that is both simple and consistent with the analysis in the rest of the paper, I include as leisure (i) activities used to construct LIU in Table 1 ("CIE," "On the phone," and "Watching TV"); (ii) activities used to construct the four leisure categories in Table 2 ("Watching TV," "Reading/Writing," "Socializing," and "Exercise"); and (iii) "Relaxing/Resting" (the only remaining activity that is unambiguously leisure). This set of activities ensures that all of the individual leisure variables constructed for Tables 1 and 2 are subsets of TLT, which is necessary for the relevance of TLT as an instrument. Using these seven activities, I then construct the TLT instrument as follows: I allocate 60 minutes to TLT if a respondent reports any two of the above-mentioned leisure activities in a particular hour,[32] 30 minutes to TLT if a respondent reports any one of the above-mentioned leisure activities and another activity (not included above) in a particular hour,[33] and zero minutes otherwise.

To allay concerns that the precise choice of leisure activities could be driving the results, in Appendix C I construct the TLT instrument using several more expansive sets of leisure activities and an alternative definition that counts leisure activities that occur in the same hour as "Other" as 30 minutes of leisure (rather than a full hour). The results are unaffected by these alternative definitions.

B.3 Summary statistics

Table A1 reports summary statistics for the constructed time-use variables described in the main text.

Table A1: Summary statistics for constructed time-use variables

	Mean	Std. Dev.	Min.	P_{25}	P_{50}	P_{75}	Max.
LIU	72.1	90.1	0	0	60	120	780
TV	107.3	99.3	0	30	90	180	840
Reading/Writing	26.3	52.0	0	0	0	30	840
Exercise	17.5	39.5	0	0	0	0	540
Socializing	32.9	71.8	0	0	0	30	840

Source: Survey of Unemployed Workers in New Jersey

Notes: All times are reported in minutes per day. Sample consists of all respondents.

[32] If a respondent reports only one of these activities and either (i) no second activity or (ii) "Other" as the second activity, I also allocate 60 minutes to TLT. The former is important but almost certainly correct, while the latter is more ambiguous but inessential for the results; see Appendix C.

[33] To be consistent with the definition of LIU in (8), I count time spent on CIE and any element of $\mathbb{NL}^{\text{Online}}$ from Section 3 (that is, work-related activities that typically occur online) as zero minutes of leisure.

C Additional results and robustness

This appendix reports various robustness checks for the results reported in Table 1 of the main text.

C.1 Full results

Table A2 reports full results (i.e., including coefficients on other covariates) for the main regression in Table 1 of the main text.

Table A2: Habit-formation in LIU

	Millennials		*Generation-X*		*Baby Boomers*	
	A-H	TLT	A-H	TLT	A-H	TLT
Habit ($\tilde{\rho}$)	-0.11	0.02	0.13***	0.37***	0.06*	0.08
	(0.14)	(0.20)	(0.05)	(0.14)	(0.04)	(0.06)
Duration	0.36	0.05	-1.34	-1.07	1.33	1.38
	(3.80)	(3.97)	(1.38)	(1.48)	(1.25)	(1.24)
Weekend	-11.78	-12.43	-1.78	-1.25	-13.85***	-13.84***
	(12.49)	(13.39)	(7.30)	(8.07)	(4.65)	(4.71)
Female	-3.32	-3.56	0.07	0.32	1.16	1.14
	(9.05)	(9.25)	(3.01)	(3.25)	(2.11)	(2.12)
Savings	-14.80*	-15.77*	2.20	0.87	0.96	0.93
	(8.17)	(8.76)	(2.59)	(2.92)	(2.35)	(2.37)
College grad.	-33.68**	-32.46**	0.84	-3.86	-2.52	-2.51
	(14.13)	(14.51)	(6.18)	(7.59)	(10.71)	(10.71)
Married	-12.18	-12.68	0.66	1.89	3.46	3.42
	(8.42)	(8.95)	(4.62)	(5.25)	(3.93)	(3.94)
Kids	1.30	3.00	-1.27	-2.44	-2.43	-2.42
	(12.07)	(13.52)	(4.98)	(5.64)	(2.67)	(2.68)
Constant	54.76**	54.45*	8.60	16.49	-16.56	-16.89
	(26.92)	(28.34)	(12.11)	(13.76)	(17.76)	(17.78)
Dummies:						
Time	×	×	×	×	×	×
Cohort	×	×	×	×	×	×
A-B test: AR(2)†	0.95	0.61	0.85	0.24	0.08	0.11
A-B test: AR(3)†	0.65	0.59	0.85	0.80	0.33	0.33
K-P test‡	0.00	0.00	0.00	0.00	0.00	0.00
Observations	784	784	4897	4897	7880	7880

†p-values; H_0: No order-i autocorrelation in residuals.

‡p-values; H_0: Model underidentified.

Robust standard errors in parentheses.

C.2 Including incomplete time diaries

Table A3 reports results for the main regression in Table 1 without imposing the Krueger and Mueller (2011) criteria on time diary selection (i.e., dropping time diary entries with fewer than 14 complete entries or fewer than four different activities reported during a given day).

Table A3: Including incomplete time diaries

	Millennials		*Generation-X*		*Baby Boomers*	
	A-H	TLT	A-H	TLT	A-H	TLT
Habit ($\tilde{\rho}$)	-0.19	-0.08	0.15***	0.22**	0.07**	0.12**
	(0.14)	(0.16)	(0.04)	(0.11)	(0.03)	(0.06)
Full controls:	×	×	×	×	×	×
A-B test: AR(2)†	0.22	0.63	0.46	0.18	0.05	0.03
A-B test: AR(3)†	0.24	0.22	0.33	0.33	0.29	0.29
K-P test‡	0.00	0.00	0.00	0.00	0.00	0.00
Observations	931	931	5437	5437	8771	8771

†p-values; H_0: No order-i autocorrelation in residuals.
‡p-values; H_0: Model underidentified.
Robust standard errors in parentheses.
Notes: Sample includes all time diaries regardless of the number of completed entries or distinct activities reported.

C.3 Excluding outlying observations

Table A4 reports results after excluding observations for which $\Delta l_{i,t}$ is below the 1^{st} percentile or above the 99^{th} percentile of the distribution of $\Delta l_{i,t}$.

Table A4: Trimming outlying observations

	Millennials		*Generation-X*		*Baby Boomers*	
	A-H	TLT	A-H	TLT	A-H	TLT
Habit ($\tilde{\rho}$)	-0.05	0.07	0.23***	0.58***	0.06*	0.08
	(0.11)	(0.17)	(0.05)	(0.18)	(0.04)	(0.08)
Full controls:	×	×	×	×	×	×
A-B test: AR(2)†	0.77	0.52	0.91	0.17	0.02	0.06
A-B test: AR(3)†	0.47	0.45	0.76	0.87	0.05	0.05
K-P test‡	0.00	0.00	0.00	0.00	0.00	0.00
Observations	754	754	4740	4740	7676	7676

†p-values; H_0: No order-i autocorrelation in residuals.
‡p-values; H_0: Model underidentified.
Robust standard errors in parentheses.
Notes: Sample excludes outliers, defined as observations below the 1^{st} percentile or above the 99^{th} percentile of the distribution of $\Delta l_{i,t}$.

C.4 Alternative definitions of LIU

Table A5 reports results for four alternative definitions of LIU. *Incl. R/W* adds "Reading/Writing" to $\mathbb{L}^{\text{Online}}$; *Incl. R/W/S* adds "Reading/Writing" and "Socializing" to $\mathbb{L}^{\text{Online}}$; *Excl. search* only includes "Job Search" in $\mathbb{NL}^{\text{Online}}$; and *Alt. CIE* counts a response of only CIE as 45 minutes of LIU.

Table A5: Alternative definitions of LIU

	Incl. R/W		*Incl. R/W/S*		*Excl. search*		*Alt. CIE*	
	A-H	TLT	A-H	TLT	A-H	TLT	A-H	TLT
Habit ($\tilde{\rho}$)	0.14***	0.36**	0.13**	0.35**	0.14***	0.40***	0.13***	0.35**
	(0.05)	(0.17)	(0.05)	(0.17)	(0.05)	(0.15)	(0.05)	(0.14)
Full controls:	×	×	×	×	×	×	×	×
A-B test: AR(2)†	0.53	0.17	0.50	0.17	0.85	0.23	0.47	0.12
A-B test: AR(3)†	0.46	0.39	0.45	0.39	0.98	0.88	0.65	0.63
K-P test‡	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Observations	4897	4897	4897	4897	4897	4897	4897	4897

†p-values; H_0: No order-i autocorrelation in residuals.
‡p-values; H_0: Model underidentified.
Robust standard errors in parentheses.
Notes: Incl. R/W: Definition of LIU includes time spent online also reading or writing; Incl. R/W/S: Definition of LIU includes time spent online also reading or writing or socializing; Excl. search: Definition of LIU only excludes CIE time also spent on job search (as opposed to excluding CIE time also spent on job search, working, taking a course and job training); Alt. CIE: Definition of LIU counts a response of only CIE as 45 minutes of LIU (rather than a full hour).

C.5 Alternative definitions of TLT

Table A6 reports results based on three alternative definitions of TLT. *Incl. shopping*: TLT includes "Shopping"; *Incl. eating*: TLT includes "Eating and drinking"; *Excl. other*: Activities occurring in the same hour as Other count as 30 minutes of leisure rather than 60.

Table A6: Alternative definitions of TLT

	Baseline	*Incl. shopping*	*Incl. eating*	*Excl. other*
Habit ($\tilde{\rho}$)	0.37***	0.33**	0.41**	0.31**
	(0.14)	(0.15)	(0.17)	(0.13)
Full controls:	×	×	×	×
A-B test: AR(2)†	0.24	0.33	0.21	0.35
A-B test: AR(3)†	0.80	0.80	0.79	0.81
K-P test‡	0.00	0.00	0.00	0.00
Observations	4,897	4,897	4,897	4,897

†p-values; H_0: No order-i autocorrelation in residuals.
‡p-values; H_0: Model underidentified.
Robust standard errors in parentheses.
Notes: Incl. shopping: Definition of TLT includes shopping; Incl. eating: Definition of TLT includes eating; Excl. other: Activities reported in the same hour as Other count for 30 minutes of leisure (rather than a full hour).

C.6 Difference GMM

Table A7 reports results from the baseline A-H specification in the main text as well as the Arellano-Bond difference GMM estimator, which uses additional lags of $l_{i,t-1}$ as instruments.

Table A7: Difference GMM

	Millennials		*Generation-X*		*Baby Boomers*	
	A-H	GMM	A-H	GMM	A-H	GMM
Habit ($\tilde{\rho}$)	-0.11	-0.10	0.13***	0.11**	0.06*	0.08**
	(0.14)	(0.13)	(0.05)	(0.05)	(0.04)	(0.03)
Full controls:	×		×		×	
Hansen test‡		0.43		0.44		0.60
Observations	784	784	4897	4920	7880	8032

‡p-values; H_0: Over-identifying restrictions valid.

Two-step standard errors in parentheses.

Notes: Results based on instrumenting with the first three lags of $l_{i,t-1}$.