I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17101

The Importance of Sampling Frequency for Estimates of Well-Being Dynamics

Stephen Hoskins
David W. Johnston
Johannes S. Kunz
Michael A. Shields
Kevin E. Staub

JUNE 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17101

The Importance of Sampling Frequency for Estimates of Well-Being Dynamics

Stephen Hoskins
Singapore Management University

David W. Johnston
Monash Business School

Johannes S. Kunz
Monash Business School

Michael A. Shields
Monash Business School and IZA

Kevin E. Staub
University of Melbourne and IZA

JUNE 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Importance of Sampling Frequency for Estimates of Well-Being Dynamics*

Using a high-frequency panel survey, we examine the sensitivity of estimated self-reported well-being (SWB) dynamics to using monthly, quarterly, and yearly data. This is an important issue if SWB is to be used to evaluate policy. Results from autoregressive models that account for individual-level heterogeneity indicate that the estimated persistence using yearly data is near zero. However, estimated persistence from monthly and quarterly data is substantial. We estimate that persistence to shocks typically lasts around six months and has a net present value of 75-80 per cent of the contemporaneous effect. Estimates are similar for different domains of SWB.

JEL Classification: I1, I3

Keywords: well-being, life satisfaction, happiness, dynamic panel data, panel autoregression, adaptation, persistence

Corresponding author:
Kevin Staub
Department of Economics
University of Melbourne
111 Barry Street
3010 VIC
Australia

E-mail: kevin.staub@unimelb.edu.au

* We thank the Center for Research on Successful Ageing at the Singapore Management University for the use of the Singapore Life Panel data. We thank The Ngee Ann Kongsi and the Ministry of Education, Singapore (under its Academic Research Fund Tier 3 program, award reference numbers MOE2013-T3-1-009 and MOE2019-T3-1-006) for financial support. Authors' contributions (CRediT): S Hoskins: Data curation. DW Johnston: Conceptualization, methodology, writing. JS Kunz: Conceptualization, methodology, formal analysis, writing. KS Staub: Conceptualization, methodology, formal analysis, writing. MA Shields: Conceptualization, methodology, writing.

1 Introduction

Self-reported well-being (SWB) data is often used to estimate changes in a persons well-being over time due to individual adversity, as well as large-scale events, and there are calls to use SWB to evaluate competing policy choices (e.g. Clark, 2018; De Neve et al., 2020; Frijters et al., 2015). However, the literature on the extent and speed of adaptation to events has mostly analysed yearly panel data (e.g. Clark et al., 2008), and there is little evidence on the sensitivity of results to how frequently SWB is measured. Importantly, shocks that appear largely contemporaneous when studied with yearly data (i.e. have low yearly persistence), may actually cause substantial within-year changes. In this study, we explore the extent and duration of SWB dynamics by estimating the monthly, quarterly and yearly persistence of SWB shocks.

Notably, we use a dynamic SWB model rather than one based on observed events. Thus, the dynamics we study subsume all shocks, giving a big-picture perspective of persistence, of which there is little evidence. A recent exception is Piper (2023), who concludes from analyses of yearly data that there is low persistence in SWB and that SWB determinants are mostly contemporaneous.[1] Our results support a different conclusion.

2 The Singapore Life Panel

Our data come from 37 consecutive waves of the Singapore Life Panel (SLP) spanning September 2015 to September 2018. The SLP is a monthly online household panel survey of about 11,000 Singaporeans aged 50 to 70 (see Vaithianathan et al., 2018, for more information on the SLP). Each month, panel members are asked questions about various topics, including their satisfaction with life, our measure of SWB. Our estimation sample is a balanced panel of 3,055 individuals with non-missing life satisfaction responses, providing 113,035 observations.

Life satisfaction is measured with the question "Taking all things together, how satisfied are you with your life as a whole these days?" to which respondents could answer on a five-point scale (very dissatisfied to very satisfied). The survey also includes satisfaction questions about social contacts and family life, daily activities and work, income, economic situation, and health.

[1] Also see Bottan and Truglia (2011) and Wunder (2012).

3 Life satisfaction dynamics with yearly, quarterly and monthly data

We first estimate persistence with a simple autoregressive model:

$$y_{i,t} = \rho y_{i,t-s} + x'_{i,t}\beta + \varepsilon_{i,t}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{1}$$

where $y_{i,t}$ denotes the life satisfaction of individual i in month t. To understand the overall persistence in life satisfaction, we include only a parsimonious set of exogenous variables in $x_{i,t}$: indicators for gender, age in years, and year of survey.

We use two approaches to examine the importance of sampling frequency on ρ. First, we include all available monthly observations and vary the lag length (monthly s=1, quarterly s=3, and yearly s=12). Second, we mimic lower-frequency panel data sets by constructing artificial panels: quarterly panel includes December, March, June, and September responses, and yearly panel includes December responses.

Table 1 presents estimates of ρ for different frequencies. The estimates are similar across columns, ranging from 0.728 (yearly panel) to 0.808 (monthly panel), indicating a high period-to-period persistence regardless of the sampling frequency. An autoregressive parameter of 0.8, for example, implies that a one-unit shock in period t will have died away (arbitrarily defined as magnitude < 0.05) within 14 years for the yearly panel and 14 months for the monthly panel. This highlights the speciousness of similar persistence estimates across different frequencies.[2]

4 Sources of persistence: unobserved heterogeneity vs state dependence

The large persistence estimates in Table 1 indicate that high life satisfaction today predicts high life satisfaction levels in the future. We next establish how much of the variation in estimates is driven by persistence across time within individuals (i.e. state dependence) and time-invariant differences between individuals (i.e. unobserved heterogeneity). To do so, we augment equation (1) with an individual-specific, time-invariant component:

[2] Selective attrition does not drive these results. Appendix Table A1 reports estimates using the unbalanced sample, and estimates are very similar to those in Table 1.

$$y_{i,t} = \rho y_{i,t-s} + x'_{i,t}\beta + \alpha_i + \varepsilon_{i,t}. \tag{2}$$

ρ now captures the average persistence in life satisfaction within an individual.

It is well-known that the fixed-effects estimator is inconsistent for linear dynamic models (Anderson and Hsiao, 1981), so we use the split-panel jackknife estimator of Dhaene and Jochmans (2015). If one considers T to be fixed, this estimator is asymptotically equivalent to the GMM estimators of Arellano and Bond (1991) and Blundell and Bond (1998). However, moderately long panels such as the SLP are better described by asymptotics in which T is not fixed. Under these conditions the jackknife estimator is asymptotically efficient (Dhaene and Jochmans, 2015).

Table 2 contains yearly, quarterly, and monthly persistence estimates from the naïve estimator (which suffers from Nickell bias) and the bias-corrected estimator.[3] Yearly persistence estimates are near-zero, indicating the Table 1 estimate was driven by time-invariant differences between individuals. In contrast, only roughly 50% of quarterly and monthly persistence are due to time-invariant differences. Thus, leveraging the higher sampling frequency makes it possible to detect substantial within-year dynamic movements in life satisfaction. We next examine this within-year adaptation process.

5 High-frequency estimates of the response to a shock

The monthly estimate (0.436) from Table 2 represents the most finely estimated autoregressive parameter from our data. We use it to trace the adaptation path of a Δ-sized shock to life satisfaction in Figure 1. The red line shows how the shock persists over the next 12 months: $\rho\Delta$ after one month, $\rho^2\Delta$ remains after two months, etc. After six months, the effect is close to zero.

Model (2) is highly stylised, with dynamics represented by a single parameter. While this has parsimony advantages, we can use a less parametric approach by specifying an AR(12) model:

$$y_{i,t} = \sum_{s=1}^{12} \rho_s y_{i,t-s} + x'_{i,t}\beta + \alpha_i + \varepsilon_{i,t}, \tag{3}$$

where the persistence is estimated flexibly via the twelve parameters (ρ_1 to ρ_{12}).

[3] Table 2 highlights the importance of Nickell biases, with bias-corrected estimates substantially larger than naïve ones. While the bias diminishes with T, simulations in Chudik, Pesaran and Yang (2018) demonstrate substantive biases at T=30 (and even some bias with T=200). Moreover, if there are other weakly exogenous regressors besides the lagged dependent variable, the naïve fixed effects estimator is misspecified and will suffer from additional biases. In contrast, the bias-corrected estimator addresses bias from all such regressors. Finally, further misspecification, such as heterogeneity in state dependence, might affect naïve and bias-corrected estimators differentially (cf. Hoskins et al., 2024).

The blue line in Figure 1 shows the estimated adaptation path from the AR(12).[4] Compared to the AR(1), the AR(12) gives a flatter adaptation trajectory. The immediate depreciation of the life satisfaction shock is larger, with only 25% left after one month. However, the shock persists for longer than the AR(1) implied, disappearing after nine months.

The simple AR(1) appears to be effective at uncovering the presence of persistence. However, the adaptation path may be different than implied by an AR(1). An important metric to assess such differences is net present value (NPV). I.e., the difference in future SWB an individual receives from an initial Δ. Using standard discount rates (e.g. Baetschmann, Staub and Studer, 2016), we find that NPV estimates are similar for both models (see Table 3), varying between 81.5% and 75.0% of the initial shock. However, the NPV for the quarterly AR(1) model is substantially smaller, with values of only 44%-48%. Therefore, the results again indicate that a higher sampling frequency helps detect within-year dynamic movements in life satisfaction.[5]

6 Dynamics in life satisfaction domains

Finally, we explore the dynamics of satisfaction with different domains of life to determine whether particular domains drive the persistence observed above. Specifically, we estimate yearly autoregressive models without fixed effects for each domain, as in equation (1), and yearly, quarterly, and monthly models with fixed effects, as in equation (2). The dynamics for each domain are similar to the dynamics of overall life satisfaction (see Table 4). The largest monthly persistence estimate occurs for income shocks (0.316), and the smallest occurs for social shocks (0.258). Thus, we find little evidence of the differential contribution of satisfaction domains to the persistence of life satisfaction.

7 Discussion

This paper demonstrates the existence of substantial short-term dynamics in life satisfaction. These dynamics add 75–80% to the net present worth of a life satisfaction shock and are also present in important domains of life.

This persistence cannot be detected by yearly frequency data that most life satisfaction research relies upon. Therefore, researchers working with such data should be aware that estimated effects on life satisfaction may overlook substantial short-term dynamics, which is an important consideration if

[4] AR(12) estimates are reported in Appendix Table A2.

[5] While the additional effect of monthly persistence is substantial, it is likely an underestimate of the NPV that could theoretically be obtained with satisfaction measured in continuous time.

SWB measures are to be used to evaluate policies. Quarterly data also significantly underestimates the degree of persistence, but it can be useful for testing the presence of persistence.

While this paper's results were obtained using a representative sample of older individuals, it is plausible to assume that short-term dynamics also exist among younger individuals. However, regardless of this, inferences about dynamics in the general population based on low-frequency data will be biased by the significant short-term dynamics among the older cohorts, given that they constitute a substantial share of the population.

Higher sampling frequency is costly relative to yearly data collection. However, the SLP (and other panels such as the American Life Panel, Pollard and Baird, 2017) show that implementing such monthly panel surveys is feasible. Moreover, there are possibilities to obtain higher frequency information at lower cost. These include limiting the higher-frequency sampling to (random or targeted) subgroups of respondents in yearly household panels and using inexpensive collection methods such as online or SMS surveys instead of in-person or over-the-phone interviews. Governments interested in collecting higher-frequency satisfaction measures could also link such collection to already existing (higher-frequency) routine administrative data collection.

References

Anderson, TW 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*.

Baetschmann, G, KE Staub and R Studer. 2016. "Does the stork deliver happiness? Parenthood and life satisfaction." *Journal of Economic Behavior and Organization*.

Blundell, R and S Bond. 1998. "Does the stork deliver happiness? Parenthood and life satisfaction." *Journal of Econometrics*.

Bottan, NL and RP Truglia. 2011. "Deconstructing the hedonic treadmill: Is happiness autoregressive?" *The Journal of Socio-Economics*.

Chudik, A, MH Pesaran and JC Yang. 2018. "Half-panel jackknife fixed-effects estimation of linear panels with weakly exogenous regressors." *Journal of Applied Econometrics*.

Clark, A, E Diener, Y Georgellis and R Lucas. 2008. "Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis." *Economic Journal*.

Clark, AE. 2018. "Four decades of the economics of happiness: Where next?" *Review of Income and Wealth.*

De Neve, JE, AE Clark, C Krekel, R Layard and G O'Donnell. 2020. "Taking a wellbeing years approach to policy choice." *BMJ* .

Dhaene, G and K Jochmans. 2015. "Split-panel jackknife estimation of fixed-effect models." *Review of Economic Studies.*

Frijters, P, AE Clark, C Krekel and R Layard. 2015. "A happy choice: Wellbeing as a goal of government." *Behavioural Public Policy.*

Hoskins, S, DW Johnston, JS Kunz, MA Shields and KE Staub. 2024. "Heterogeneity in the Persistence of Health: Evidence from a Monthly Micro Panel." *IZA Discussion Paper 17023.*

Piper, A. 2023. "What does dynamic panel analysis tell us about life satisfaction?" *Review of Income and Wealth.*

Pollard, MS and MD Baird. 2017. "The RAND American Life Panel: Technical Description." *Santa Monica, CA: RAND Corporation.*

Vaithianathan, R, B Hool, MD Hurd and S Rohwedder. 2018. "High frequency internet survey of a probability sample of older Singaporeans: The Singapore Life Panel." *Singapore Economic Review.*

Wunder, C. 2012. "Does subjective well-being dynamically adjust to circumstances?" *Economics Letters.*

Tables and figures

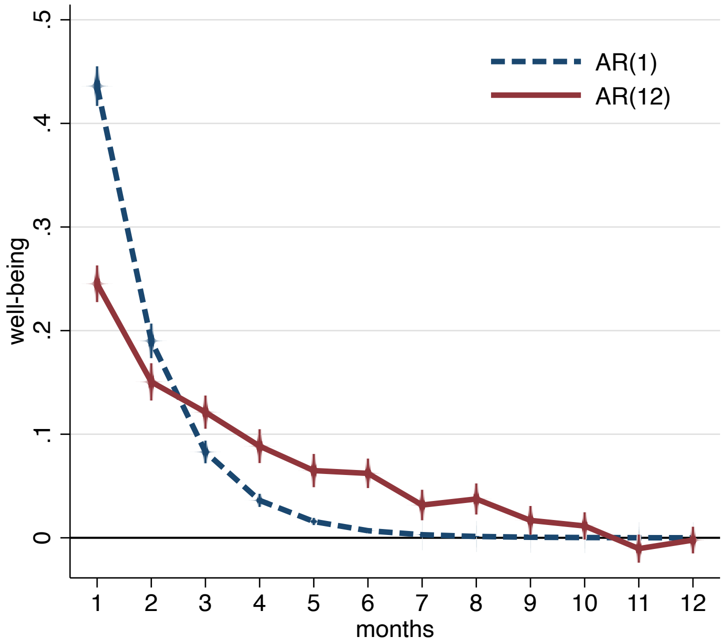

Figure 1: Impulse response plot of two dynamic life satisfaction models with monthly data

Note: Blue line: AR(12) model (equation 3) estimated by split-panel jackknife. Red line: AR(1) model from Table 2 Column (6) where $\hat{\rho}$, $\hat{\rho}^2$, $\hat{\rho}^3, \ldots$ are estimates for months 1, 2, 3,... Control variables: see Table 1 notes. Vertical lines: 99% confidence intervals from cluster-robust standard errors. See also Appendix Table A2.

Table 1: Dynamic SWB models: Pooled OLS

	(1) yearly	(2) yearly	(3) quarterly	(4) quarterly	(5) monthly
lagged life satisfaction	0.728***	0.771***	0.786***	0.792***	0.808***
	(0.009)	(0.011)	(0.006)	(0.007)	(0.005)
N	76,375	6,110	103,870	33,605	109,980

Notes: Estimates of ρ from AR(1) models. Columns 2 and 4 represent artificial panels that mimic lower-frequency panel data sets. Control variables: complete sets of year and age (in years) fixed effects, gender dummy. Robust standard errors clustered at individual level in parentheses.
Statistical significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Dynamic SWB models: FE and bias-corrected FE

	(1) yearly	(2) yearly bc	(3) quarterly	(4) quart., bc	(5) monthly	(6) monthly, bc
lagged life satisfaction	0.003	0.004	0.195***	0.328***	0.301***	0.436***
	(0.006)	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)
N	76375	76375	103870	103870	109980	109980

Notes: Within estimates (columns 1, 3, 5) and split-panel jackknife estimates (2, 4, 6) of AR(1) models with individual-specific fixed effects. Additional information: see Table 1 notes.

Table 3: Net present value of SWB shock

	Annual discount factor (κ)			
	κ=1	κ=0.95	κ=0.90	κ=0.80
monthly AR(12) model, fe bc	0.815	0.804	0.794	0.773
monthly AR(1) model, fe bc	0.773	0.767	0.762	0.750
quarterly AR(1) model, fe bc	0.482	0.474	0.465	0.448

Notes: Calculations based on Cols (2) [monthly AR(12)] and (3) [monthly AR(1)] of Table A2, and (monthly prorated) annual discount factors κ. Quarterly AR(1) model estimates based on Table 2 Col (4) and quarterly prorated factors.

Table 4: Dynamic SWB models: domain satisfactions

	(1) pooled year	(2) bc fe year	(3) bc fe quart	(4) bc fe month
lagged satisfaction with family	0.727***	0.004	0.167***	0.258***
	(0.009)	(0.005)	(0.007)	(0.007)
N	76332	76332	103816	109921
lagged satisfaction with income	0.772***	-0.008	0.208***	0.316***
	(0.007)	(0.006)	(0.008)	(0.008)
N	76322	76322	103803	109908
lagged satisfaction w. econ. situation	0.764***	-0.010	0.173***	0.281***
	(0.007)	(0.005)	(0.007)	(0.007)
N	76315	76315	103786	109890
lagged satisfaction with health	0.762***	0.000	0.158***	0.277***
	(0.008)	(0.005)	(0.007)	(0.007)
N	76341	76341	103822	109930
lagged satisfaction with job	0.710***	-0.001	0.166***	0.268***
	(0.009)	(0.005)	(0.007)	(0.007)
N	76288	76288	103760	109865

Notes: Each panel depicts a separate AR(1) model. Cols. (1)-(2) include a yearly lag, (3) a quarterly lag, (4) a monthly lag. Additional information: see Table 1 notes.

Additional tables for online appendix

Table A1: Dynamic SWB models: Pooled OLS, unbalanced samples

	(1) yearly	(2) yearly	(3) quarterly	(4) quarterly	(5) monthly
lagged life satisfaction	0.689***	0.729***	0.755***	0.759***	0.785***
	(0.006)	(0.008)	(0.004)	(0.005)	(0.004)
N	184332	15246	251362	80332	263563

Notes: Pooled OLS estimates of the autoregressive coefficient in AR(1) models. of Control variables, standard errors and significance: see Table 1 notes.

Table A2: Dynamic SWB models: AR(12) models with monthly data

	(1) fe	(2) fe, bc	(3) fe, bc, based on ar(1)
Life satisfaction, 1st lag	0.156***	0.245***	0.436***
	(0.007)	(0.007)	(0.007)
Life satisfaction, 2nd lag	0.078***	0.150***	0.190***
	(0.007)	(0.007)	(0.006)
Life satisfaction, 3rd lag	0.059***	0.121***	0.083***
	(0.006)	(0.006)	(0.004)
Life satisfaction, 4th lag	0.035***	0.088***	0.036***
	(0.006)	(0.006)	(0.002)
Life satisfaction, 5th lag	0.020**	0.065***	0.016***
	(0.006)	(0.006)	(0.001)
Life satisfaction, 6th lag	0.024***	0.063***	0.007***
	(0.006)	(0.006)	(0.001)
Life satisfaction, 7th lag	0.000	0.031***	0.003***
	(0.006)	(0.006)	(0.000)
Life satisfaction, 8th lag	0.012*	0.037***	0.001***
	(0.006)	(0.006)	(0.000)
Life satisfaction, 9th lag	-0.003	0.017**	0.001***
	(0.005)	(0.005)	(0.000)
Life satisfaction, 10th lag	-0.000	0.011*	0.000***
	(0.005)	(0.005)	(0.000)
Life satisfaction, 11th lag	-0.015**	-0.011*	0.000***
	(0.005)	(0.005)	(0.000)
Life satisfaction, 12th lag	0.000	-0.002	0.000***
	(0.005)	(0.005)	(0.000)
N	76375	76375	76375

Note: Estimates of the autoregressive coefficients in AR(12) models (Cols. 1–2) and AR(1) models (Col. 3). Col. (1) obtained using the within estimator, Cols. (2) and (3) using the split-panel jackknife estimator. Estimates of the autoregressive coefficients of order higher than 1 in Col. (3) obtained as powers of $\hat{\rho}$: $\hat{\rho}^2$, $\hat{\rho}^3$, etc. for second lag, third lag, etc. Standard errors for Col. (3) based on Delta method. Estimates from Cols. (2) and (3) are plotted in Fig. 1. Control variables, standard errors and significance: see Table 1 notes.