

Evidence for Sticky Consumption Growth in Canada

by Christopher J. Lavoie

6874698

Major Paper presented to the
Department of Economics of the University of Ottawa in partial
fulfillment of the requirements of the M.A. Degree
Supervisor: Professor Maral Kichian

ECO 6999

Ottawa, Ontario
August 2020

Abstract

Consumption stickiness occurs when current consumption can be predicted by past consumption. The presence of consumption stickiness has important implications on our understanding of consumption dynamics and how they respond to shocks in income. I use the instrument-based framework by Carroll, Slacalek, and Sommer (2011) for the case of Canada over the period 1980Q1 to 2020Q1. I show strong evidence for the presence of weak instruments in the data and therefore use weak-instrument robust inference techniques: the conditional likelihood ratio test (Moreira, 2003) and the Anderson Rubin test (Anderson & Rubin, 1949). I find strong evidence of consumption stickiness in the case of Canada for all examined subsamples.

Contents

1	Introduction	3
2	Consumption Stickiness: Literature Review	5
2.1	Sticky Consumption Literature	5
2.2	Sticky Consumption Models	8
3	Instrumental Variable Analysis	10
3.1	Model	11
3.2	Instruments used in consumption stickiness	12
3.3	Weak Instruments	12
3.4	Anderson-Rubin Estimates	13
3.5	Conditional Likelihood Ratio Test	14
4	Testing for Weak Instruments	14
5	Estimates of Consumption Stickiness	15
5.1	Main Specification	15
5.2	Data	15
5.3	Main Results	17
5.4	Robustness Checks	20
5.5	Stationarity of Main Specification	22
6	Discussion	24
6.1	Macroeconomic Considerations	24
6.2	Econometric Considerations	27
7	Conclusion	30
	References	32

1 Introduction

Since Hall's (1978) work hypothesizing aggregate consumption dynamics as a random walk with trend, much work has been devoted to empirically analyzing aggregate consumption dynamics. The interest in this problem has a clear motivation: understanding how consumption changes based on changes in income is crucial for understanding aggregate responses to economic shocks. Whether through a large economic crash or through stimulus checks that boost aggregate income, understanding how consumption relative to that income changes over time is essential to determining the possible effects of these shocks.

However, understanding the dynamics of consumption growth is not so straightforward. Campbell and Deaton (1989) found that the behaviour of aggregate consumption relative to aggregate income is excessively smooth compared to Hall's permanent income hypothesis. Instead of consumers instantly adjusting to changes in income as predicted by Hall, Fuhrer (2000) finds that economic variables exhibit a gradual response that occurs over time. Fuhrer notes that this gradual response looks like a hump-shaped curve, and uses a model of habit formation, where utility depends on both current and past consumption, to estimate this effect. Since that study, this behaviour in consumption has been known as consumption stickiness, and can be characterised econometrically by a "stickiness" parameter that estimates the impact of past consumption on current consumption (Dynan, 2000; Carroll et al., 2011). Fuhrer (2000) finds that this stickiness parameter determines the degree to which the response is hump-shaped; a higher parameter indicates a higher hump-shaped response and therefore exhibits higher excess smoothness.

While Fuhrer (2000) finds evidence for habit formation (and therefore sticky consumption) in aggregate data, Dynan (2000), finds no evidence of habit formation when testing on annual data on food consumption in the US. Such disparities in results are not unusual in the literature. Empirical estimates of habit formation are highly dependent on the type of data used, the frequency of the data, and the goods that are being measured (Havranek, Rusnak, & Sokolova, 2017).

One important work in this literature is from Carroll et al. (2011). Carroll et al. use aggregate quarterly data for 13 advanced economies, including Canada, and find a relatively similar level of consumption stickiness of approximately 0.7 in all countries during the period from 1962Q3-2004Q2, although the specific time period used for each country varies based on the availability of data. In the case for Canada the time period of Carroll et al.'s estimates is from 1970Q4-2002Q3. Since this study, there have been new developments in the macroeconomic environment that may warrant an updated analysis. The data after 2002Q3 contain

observations on macroeconomic variables during the Financial Crisis of 2008-2009, which caused one of the most significant economic disruptions to advanced economies worldwide since the Great Depression in the 1930s. Kumar and Jia (2019) have found that estimates of consumption stickiness in the US fall sharply in the wake of the financial crisis, and recover back to levels consistent with those estimated by Carroll et al. (2011) afterward. Moreover, the extent of stickiness may vary over different subperiods, depending on economic conditions or the evolution of structural macroeconomic parameters over the past fifty years.

In this Major Research Paper, I estimate consumption stickiness for the Canadian case using the framework from Carroll et al. (2011) with aggregate data and instrument-based methods. I first report evidence for weak instrument sets both for the original Carroll et al. (2011) estimated model and in my own preferred specification with updated data. Given the presence of weak instruments, I employ the Conditional Likelihood Ratio method to test for significance of my estimates in a way that is robust to weak instruments, as in Carroll et al. (2011). I thus provide the estimates of the consumption stickiness parameter by Carroll et al. over the period of 1980Q1-2002Q3, and then extend the analysis with data up to 2020Q1. To determine if the financial crisis had a significant impact on consumption stickiness estimates in Canada, as found in Kumar and Jia (2019) for the US, I also obtain estimates of consumption stickiness over rolling subsamples. In these rolling subsamples, I employ the weak-instrument robust Anderson Rubin test method, which is also robust to excluded instruments, in addition to the Conditional Likelihood Ratio test. Finally, I verify the results under a series of robustness checks.

Overall, I find evidence for consumption stickiness across the entire sample from 1980Q1 to 2020Q1, with point estimates that on average are similar to those found in Carroll et al. (2011), however the weak-instrument robust confidence intervals are wide, revealing the lack of precision of these estimates. In addition, I also estimate the stickiness parameter over subsamples and I find that in periods where there have been significant shocks to the Canadian economy, the stickiness parameter drops, which is consistent with Kumar and Jia (2019), although in robustness checks the effect is less pronounced. Using rolling estimates, I also observe that statistical significance drops after 2016, which is likely attributed to the instrument set fitting more weakly after the 2008 financial crisis. The sensitivity analysis finds that statistical significance is robust to seasonal adjustments in the instruments, first differences in instruments, and different confidence indicators. In addition, the levels of the parameter estimates change as well.

These findings are an important contribution to the literature, as they shed light on the strength of commonly employed instrument sets throughout different time horizons over the past 40 years. Section 2

reviews the literature of consumption stickiness and some of the key models used to estimate it. Section 3 contains a brief review of weak instruments, how it complicates inference under conventional the Two Stage Least Squares (TSLS) estimation method, and methods of obtaining valid confidence intervals in the presence of weak instruments. Section 4 tests for weak instruments in Carroll et al. (2011), Section 5 presents the model, the description of the data and the main results of the study, section 6 discusses the macroeconomic and econometric considerations of the study and section 7 concludes.

2 Consumption Stickiness: Literature Review

2.1 Sticky Consumption Literature

Consumption dynamics are an important piece of the empirical macroeconomic literature. Simple models of the economy picture consumers responding instantaneously to economic factors, rationally taking into account expectations of lifetime income into consumption. One of the key features of these models is the notion of time separability of utility, where utility of consumption in one period is independent of consumption in other periods. However behavioural peculiarities in individuals may not necessarily follow time-separability. Muellbauer (1988) for example, presents a model of habit formation where utility from consumption is derived relative to the previous period. Studies like those from Fuhrer (2000) and Dynan (2000) investigate this concept econometrically. Fuhrer (2000) has demonstrated that the incorporation of habit formation helps to fit real-world aggregate consumption data, which reacts smoothly and in a hump-shaped manner in response to shocks in income. Dynan (2000) however, using household-level data on food consumption and non-durable goods and services, rejects the hypothesis of habit formation. This idea of habit formation has been elaborated on in the macro and micro literature, with estimates of habit formation differing in frequency, aggregation and definitions of consumption (Havranek et al., 2017).

The presence of consumption stickiness in aggregate data and the lack thereof in individual data is a puzzle that has spurred the generation of several alternative explanations to habit formation. Luo et al. (2017) design a general equilibrium model in which agents have limited information due to a finite capacity for information processing. The general intuition behind the model is that information about random shocks in the income process occurs throughout the consumer's lifetime, and they must choose a noisy information channel to reduce uncertainty about the innovations in income, though they are subject to a finite information processing capacity (Luo et al., 2017). In a partial equilibrium version of the permanent income hypothesis with rational inattention, Luo (2008) shows that habit formation and rational inattention

can be observationally equivalent under aggregate data, but not under individual data, since the endogenous noise faced by consumers with rational inattention would increase volatility in the individual observation, but would be eliminated in aggregation. Luo et al. (2017) show that in a general equilibrium setting of habit formation with full information and a rational inattention model with no habits, the habit formation model induces a higher equilibrium interest rate, whereas the rational inattention model has lower interest rates.

Carroll et al. (2020) present a sticky expectations model where consumers are generally inattentive, and a constant fraction of consumers update their information about future wealth periodically. Those who update their information behave as if they had full rational expectations, however, those who do not update their information carry their outdated expectations into the next period, believing them to be true. Carroll et al. (2020) find that this phenomenon leads to sticky consumption in aggregate, since only a fraction of consumers will be attentive to shocks in aggregate income, and the rest will lag in their response, but at the individual level, inattentiveness to aggregate shocks is not perceptible, as these shocks are simply aggregations of white-noise innovations. The authors conduct simulations using both aggregate and microeconomic data, and find that stickiness in consumption is present in aggregate data, however not in the microeconomic counterpart, which is consistent with the average estimates in the literature (Havranek et al., 2017).

Pagel (2017) introduces a model where consumers earn utility through two components: consumption utility and a gain-loss utility relative to a reference level, where consumption less than the reference level is more painful than consumption above the level is enjoyable (Pagel, 2017). The consumer maximizes this utility by pure consumption, their consumption relative to their beliefs about consumption in the prior period, and lifetime expectations of consumption subject to their budget constraint. The consumer experiences a vector of random shocks to both permanent and transitory income over their life cycle and they construct their beliefs about future income with rational expectations. Pagel finds that the consumer experiences excess smoothness and excess sensitivity, since the consumer prefers to spread out reductions in consumption over time in the event of adverse income shocks and prefers to allocate gains in consumption from a positive income shock over time to protect against future uncertain and disproportionately painful losses.

The notion of using behavioural peculiarities in expectations to explain dependence on lagged macroeconomic variables has some empirical support. Fuhrer (2017) uses surveys of expectations from the Survey of Professional Forecasters for inflation and unemployment rates (where unemployment rates serve as proxies for expectations of output) as data to be used in the Euler equations of an otherwise standard DSGE model. By replacing rational expectations with real survey data, Fuhrer (2017) finds the model is better specified

and depends much less on the use of lagged variables for a better empirical fit. While not directly related to consumption dynamics, this work demonstrates the ability for deviations from rational expectations to explain excess smoothness rather than modifications to standard utility functions.

Carroll et al. (2011) demonstrate that stickiness in consumption, be it explained through habit formation or inattentiveness of consumers to new information, exists in the thirteen advanced economies they examine. They use both an Instrumental Variables approach and a Kalman Filtering technique to find stickiness in consumption growth across these countries using quarterly aggregate data. Furthermore, they find that this sticky consumption model fits the data better than the random walk hypothesis of consumption growth or the Campbell and Mankiw (1989) model of hand-to-mouth consumers. Kichian and Mihic (2018) use identification robust estimation techniques with a calibrated stickiness parameter to estimate the impact of housing and financial wealth on Canadian consumption, and find indirect support for a high value for the stickiness parameter (0.8). Another approach used by both Galli (2019) and Everaert et al. (2017) is to estimate consumption stickiness using Bayesian estimation techniques, which allows them to examine a distribution of wealth effects. Galli (2019) examines consumption stickiness in Switzerland between 1981-2002 using annual consumption data, and finds that the posterior median of consumption stickiness is between 0.41-0.6. Galli also uses the Kalman Filter approach, which yields an estimate of 0.4. Everaert et al. (2017) use quarterly data from 1953-2014 in the U.S. to study excess sensitivity in aggregate consumption. They control for consumption stickiness, which was shown to affect estimates of excess sensitivity, and find a stickiness coefficient of 0.53. Everaert and Pozzi (2014) use yearly aggregated data for 15 OECD countries from 1972-2007 to estimate consumption stickiness while also allowing for variable interest rates, hours worked, income and government consumption and find little evidence for stickiness in consumption growth, although they find the impact of aggregate disposable income growth on aggregate consumption growth is positive and statistically significant.

Estimation of sticky consumption can also be done by focusing on subsamples and comparing the results. Kumar and Jia (2019) use an approach similar to Carroll et al. (2011), however, they estimate whether the degree of stickiness changes in response to economic downturns. Kumar and Jia focus on the financial crisis of 2008-2009, conducting sequential estimation of the stickiness parameter in samples approaching and including the financial crisis, and find that the degree of stickiness decreases significantly in the US from 0.73 to 0.4 in the midst of the financial crisis, but recovers back to previous levels soon afterward.

Estimation of consumption stickiness is sensitive to the details embedded within the studies. In their

meta-analysis, Havranek et al. (2017) have identified a divergence in habit formation parameter estimates depending on the usage of micro or macro data, differences in frequency of data, and the order of approximation of the Euler equation. Estimates of consumption stickiness in microeconomic data range around an average coefficient of stickiness at 0.1, with the median at 0.0, whereas macroeconomic data estimates are on average 0.57 with the median at 0.66 (Havranek et al., 2017). They also find cross-country heterogeneity, which contradicts the findings from Carroll et al. (2011). The differences in consumption stickiness between microeconomic and macroeconomic datasets are hypothesized to be due to the relatively shorter time span in microeconomic panel data or cross-sectional data compared to macroeconomic studies (Havranek et al., 2017). This is further supported by the differences in frequency of estimates. As Havranek et al. (2017) note, studies with monthly frequency result in lower stickiness estimates than those using quarterly data, and studies with annual frequency report higher consumption stickiness due to goods displaying more durability in the short run than in the longer run. However, studies that span long time periods might also run into methodological issues. In a study of the relationship of consumption growth and interest rates using both aggregate and household data, Attanasio and Weber (1993) find that differences between microeconomic cohorts estimates differ largely from macroeconomic estimates, which they attribute to demographic characteristics and to non-linearities in the aggregate figures.

2.2 Sticky Consumption Models

A basic specification used as an empirical basis for sticky consumption is the model of habit formation. The following derivation and notation is from Galli (2019). A habit-forming consumer can be modeled as having the following utility function:

$$\max_{C_t} \sum_{t=0}^{\infty} B^t U[C_t - \chi C_{t-1}] = \max_{C_t} \sum_{t=0}^{\infty} B^t U[(1 - \chi)C_t + \chi \Delta C_t] \quad (1)$$

Subject to $A_{t+1} = (1 + r)(A_t + Y_t - C_t)$. In this formulation, B , is a discount factor U is the utility function, C_t is an aggregate consumption good at time t , χ is the habit formation parameter, A_t and Y_t are total wealth and income at time t respectively, and r is the constant real interest rate.

Assuming a Constant Relative Risk Aversion utility function ($U(c) = \frac{c^{1-\sigma}}{1-\sigma}$) and a consumer with rational expectations, Galli shows the maximization problem can be represented as:

$$\Delta \log(C_{t+1} - \chi C_t) = \frac{1}{\sigma} [\log(B) + \log(1 + r)] - \frac{1}{\sigma} \log(1 + e_{t+1})$$

where e_{t+1} is the mean-zero forecast error of the consumer with rational expectations. When t is large and r is constant, the equation above can be approximated as:

$$\Delta \log C_t = \mu + \chi \Delta \log C_{t-1} + \epsilon_t \quad (2)$$

Equation 2 is useful, since it implies that consumption stickiness can be estimated econometrically by an AR(1) equation.

The interpretation of the utility function in 1 can be seen when constraining the utility to a single period:

$$U = U[(1 - \chi)C_t + \chi \Delta C_t] \quad (3)$$

In this formulation, consumers derive utility from both current levels in consumption and change in consumption when χ is between 0 and 1 (Carroll et al., 2011). When χ is 1, utility is entirely derived from changes in consumption; when χ is 0, the utility function is time-separable and becomes identical to the Hall (1978) model (Carroll et al., 2011). The χ parameter acts as the weight for a weighted average of utility derived from levels and changes in consumption. In this formulation, a parameter of χ that is greater than 1 implies that the level of consumption would contribute negatively to utility and utility is earned by the change in consumption. When χ is negative, growth in consumption decreases utility. Because of these reasons, χ is bound between 0 and 1, which is consistent with Havranek et al. (2017).

Carroll et al. (2011) provide an alternative explanation that also justifies modelling consumption growth as an AR(1) process by using a Sticky Expectations model: Consumers are inattentive to macrodevelopments. In general, a portion of consumers, who maximize their consumption as above, do not update their expectations for the future and instead assume that their prior beliefs of future wealth came true (Carroll et al., 2020). The rest update their information and proceed to maximize as if in an environment with frictionless access to information. As is shown in Galli (2019), aggregate consumption can be represented as the weighted average between these consumers:

$$\Delta C_t = \Pi \Delta C_{t-1} + (1 - \Pi) \Delta C_t^{updaters} \quad (4)$$

Where C_t is aggregate consumption at t , Π is the fraction of consumers with sticky expectations and $(1 - \Pi)$ is the fraction of updaters (Galli, 2019). $(1 - \Pi)C_t^{updaters}$ denotes the consumption of those who update their information at time t . Carroll et al. (2020) show that $(1 - \Pi)C_t^{updaters}$ is an iid variable with

mean zero. When utility is a CRRA function, equation 4 can be expressed as in equation 2 (Carroll et al., 2011).

Estimates of consumption stickiness can also come from more complex models. Everaert and Pozzi (2014) employ a model with consumption stickiness and a CRRA utility function that also includes hours worked and government consumption. Using their derivation and adapting their notation, utility can be described as follows:

$$u(C_t) = \frac{1}{1 - (1/\theta)} \left[C_t C_{t-1}^{-\chi} H_t^{-\gamma} G_t^{-\xi} \right]^{1 - (1/\theta)} \quad (5)$$

Here, $\theta \in (0, 1)$ denotes the elasticity of intertemporal substitution H_t is the total number of hours worked per capita, G_t is real per capita government consumption. χ is the habit formation parameter, ξ and γ measure the impact of government spending and per capita hours worked respectively on the marginal utility of private consumption. Everaert and Pozzi (2014) derive the first order condition for the consumers as:

$$u'(C_{t-1}) = \left(\frac{1 + r_t}{1 + \delta} \right) E_{t-1} \left[u'(C_t) \right] \quad (6)$$

Where r_t represents a variable risk-free interest rate, δ a constant rate of time preference, E_{t-1} the expectation conditional on period $t - 1$. They also note that $E_{t-1}(r_t) = r_t$. Assuming a joint normal distribution between $\Delta \ln H_t$, $\Delta \ln C_t$ and $\Delta \ln G_t$ and allowing for a proportion λ consumers to act according to Campbell and Mankiw's (1989) rule of thumb consumers, Everaert and Pozzi (2014) derive their estimable equation to be:

$$\Delta \ln C_t = a_0 + a_1 \Delta \ln C_{t-1} + a_2 \Delta \ln H_t + a_3 \Delta \ln G_t + a_4 r_t + a_5 \Delta \ln Y_t + u_t \quad (7)$$

Where a_i denotes a scalar multiple of the respective parameters from equation 5, however the parameters χ , γ , ξ , and λ are uniquely identified. Note that this model is equivalent to equation 2 under the assumption that $a_2 = a_3 = a_4 = a_5 = 0$ (Everaert & Pozzi, 2014).

3 Instrumental Variable Analysis

As is mentioned by Sommer (2007), aggregate consumption can be subject to measurement error and transitory fluctuations. Sommer remarks in the U.S. case that measurements of aggregate consumption

can be subject to sampling and non-sampling error, and that these errors can proliferate into relatively large distortions in the measurements. Sommer also notes that transitory fluctuations due to phenomena like severe weather can possibly bias these results downward. Due to complications such as these, OLS estimation of equation 2 would yield biased results. To address these issues of potential bias, I use a set of instrumental variables to obtain an unbiased estimate. However it is possible that the instruments sets I employ may be potentially weak. Below is a brief review of instrumental variable analysis in the presence of potentially weak instruments.

3.1 Model

The following model and notation is taken from Stock, Wright, and Yogo (2002) and supplemented from insights by Greene (2018).

Suppose there is the following model:

$$y = X\beta + u \tag{8}$$

$$X = Z\Pi + v \tag{9}$$

where y and X are $n \times 1$ vectors of observations on endogenous variables and Z is a $n \times K$ matrix of instruments. u and v are $n \times 1$ error terms. $[u_i v_i]'$ are assumed to be iid as $N(0, \Sigma)$. In this model the instruments are fixed and the errors are iid, and we are interested in recovering the scalar β .

In this model, $E(X_i u_i) \neq 0$, however $E(Z_i u_i) = 0$, which implies that Z is a matrix of exogenous variables. The instruments are also relevant if $\Pi \neq 0$.

The model above can be estimated via Two-Stage Least Squares (TSLS). By replacing X in equation 8 with \hat{X} where $\hat{X} = Z(Z'Z)^{-1}Z'X \equiv P_Z X$, we get:

$$\hat{\beta}_{IV} = (\hat{X}'\hat{X})^{-1}\hat{X}'y \tag{10}$$

$$= (X'P_Z'P_Z X)^{-1}X'P_Z'Y \tag{11}$$

$$= (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y \tag{12}$$

which can also be expressed as $\hat{\beta} = \frac{X'P_Z y}{X'P_Z X}$ since we have a single endogenous regressor (X is $n \times 1$). $\hat{\beta}$ is asymptotically normally distributed with mean β and a finite variance (Greene, 2018), which enables

inference using $\hat{\beta}_{IV}$ when X is endogenous.

3.2 Instruments used in consumption stickiness

In the literature there are some instruments commonly used across multiple studies to estimate consumption stickiness. Carroll et al. (2011) use unemployment rates, long-term interest rates, an index of price volatility and, where available, an index of consumer confidence as instruments for lagged consumption for 13 different countries. Kumar and Jia (2019) employ lagged unemployment rates and price volatility indexes as instruments for estimating equation 2. Galli (2019) employs a consumer sentiment index and lagged disposable income for the case of Switzerland. Everaert and Pozzi (2014) use lags of their dependent variables (hours worked, government consumption) and an error correction term that is the difference between yearly income and consumption in their preferred specifications.

3.3 Weak Instruments

In their review, Stock et al. (2002) define a concentration parameter, $\mu^2 = \Pi'Z'Z\Pi/\sigma_v^2$, which is a unitless measure for the strength of instruments within the model. μ^2 can be interpreted as the F statistic for the hypothesis that $\Pi = 0$ and they show that:

$$\mu(\hat{\beta}^{TSLs} - \beta) = (\sigma_u/\sigma_v) \frac{z_u + S_{uv}/\mu}{1 + 2z_v/\mu + S_{vv}/\mu^2} \quad (13)$$

where $z_x = (\Pi'Z'x)/(\sigma_x\sqrt{\Pi'Z'Z\Pi})$, and $S_{xy} = (y'P_zx)/(\sigma_y\sigma_x)$. In order for the normal approximation to the distribution of the TSLS estimator to be an accurate representation, μ^2 must be sufficiently large. Stock et al. (2002, p.521) define a set of instruments as weak if μ^2/K is small enough that “inferences based on conventional normal approximating are misleading.”

A common rule of thumb is a value greater than 10 for the F-statistic in the first stage of a TSLS regression, which represents the relative bias of the TSLS estimator to the inconsistency of OLS (Stock et al., 2002). However Andrews, Stock, and Sun (2019) note in their review that the rule-of-thumb statistic relies heavily on the assumption of homoscedasticity. A more fitting and general test statistic for the first stage regression estimator is the estimator from Montiel Olea and Pflueger (2013), which is robust to heteroscedasticity, autocorrelation and clustering.

To illustrate the robust test statistic, we relax the assumption of iid errors. Suppose the following assumptions from Montiel Olea and Pflueger (2013):

$$[u \ v]'[u \ v]/n \xrightarrow{p} \Omega$$

where Ω is a positive definite matrix, and

$$\begin{pmatrix} Z'u/\sqrt{n} \\ Z'v/\sqrt{n} \end{pmatrix} \xrightarrow{d} N_{2K}(0, W)$$

and

$$W = \begin{pmatrix} W_1 & W_{12} \\ W'_{12} & W_2 \end{pmatrix}$$

which is a generalized form of the previous model.

The Montiel Olea and Pflueger effective F statistic is:

$$F_{eff} = \frac{1}{n} \frac{X'Z'Z'X}{tr(\hat{W}_2)} \quad (14)$$

which is distributed as a weighted average of noncentral χ^2 variables (Andrews et al., 2019). When F_{eff} exceeds a critical value defined in Montiel Olea and Pflueger (2013), the null hypothesis of weak instruments is rejected.

3.4 Anderson-Rubin Estimates

While traditional confidence intervals from TSLS should not be used when instruments are weak, it is possible to obtain unbiased estimates and confidence sets for variables of interest with other methods. One popular method for deriving identification-robust estimates is the Anderson Rubin (1949) test statistic (henceforth, AR statistic). The statistic is defined in Stock et al. (2002) for a given β_0 as:

$$AR(\beta_0) = \frac{(y - X\beta_0)'P_Z(y - X\beta_0)/K}{(y - X\beta_0)'M_Z(y - X\beta_0)/(n - K)} \quad (15)$$

Where $M_Z \equiv I - P_Z$. Under weak instrument asymptotics, the AR statistic is distributed as χ^2_k/K regardless of the magnitude of the concentration parameter above (Stock et al., 2002). Confidence sets for the AR statistic can be derived by inverting the χ^2 test to produce a confidence interval for a given β_0 (Stock et al., 2002). The AR statistic is robust to weak instruments, excluded instruments and specification of the model for Y (Dufour, 2003). However one important drawback is that calculating this statistic also requires a full specification of the vector β_0 , which can be a practical problem when working with multiple parameters (Dufour & Taamouti, 2005).

3.5 Conditional Likelihood Ratio Test

An alternative approach developed by Moreira (2003) is a test that eliminates the dependence on the first stage regression parameter Π through conditioning. When Ω is known, one can develop a test statistic that follows a nonstandard distribution (Moreira, 2003). The Conditional Likelihood Ratio test (henceforth, CLR) has very good power properties in the homoscedastic case with a single endogenous regressor, however when errors are heteroscedastic, there is not a clear consensus on which procedures to use in the case of weak instruments (Andrews et al., 2019).

Both the CLR and AR tests have important drawbacks. Whereas the AR statistic is inefficient when the model is overidentified (Andrews et al., 2019), the CLR assumption that Ω is known can be unrealistic in applied studies, and is not robust to a situation of excluded instruments (Dufour, 2003).

4 Testing for Weak Instruments

Carroll et al. (2011) employ the CLR test for their point estimates of consumption stickiness, which is robust to weak instruments, however they do not explicitly test for weak instruments in their specification. As is discussed in section 3, the presence of weak instruments requires different statistical tests for inference compared to traditional inference. In order to compare instrument strength across specifications of the models in this paper and the original model, I conduct a first stage F test of the estimates used in Carroll et al. (2011)¹. Carroll et al. (2011) employ TSLS to estimate the stickiness parameter in consumption using the AR(1) consumption equation:

$$\Delta \log C_t = \mu + \chi E_{t-2}[\Delta \log C_{t-1}] + \epsilon_t \quad (16)$$

where C denotes consumption. The variable of interest in this regression is χ , the stickiness parameter. For the case of Canada, consumption was measured by the sum of expenditures on nondurable goods and services. Consumption in the right-hand side is instrumented by unemployment rate, long-term interest rate, an index of consumer confidence, and an index of price volatility, all lagged by 2, 3 and 4 periods.

To test for the strength of instruments, Andrews et al. (2019) recommend using a heteroscedasticity and autocorrelation robust effective F statistic, F_{EFF} from Montiel Olea and Pflueger (2013). I apply this test (as implemented by Pflueger and Wang (2015)) to the original Carroll et al. data. The results are reported

¹Carroll et al. (2011) made available the exact code and data used for the estimations in their paper. I was therefore able to apply the effective F test against their exact data to obtain a reliable estimate.

Table 1: F and Effective F test of original Carroll et al. (2011) specification for Canada.

Statistic	F	F _{EFF}
Value	3.383	4.094
Critical Value	11.52	16.014

in Table 1. These show that the null hypothesis of a weak instrument set, both under ideal conditions of homoscedasticity and in a HAC-robust scenario, cannot be rejected, implying that traditional inference using t-tests are invalid. Although Carroll et al. (2011) did not formally check for the strength of their instruments, they suspected it to be the case and thus employed the weak-instrument-robust CLR test from Andrews, Moreira, and Stock (2006).

5 Estimates of Consumption Stickiness

5.1 Main Specification

The main specification of the model to be estimated is the following:

$$\Delta \ln C_t = \mu + \chi \Delta \ln C_{t-1} + \epsilon_t \quad (17)$$

$$\Delta \ln C_{t-1} = a_0 + a_1 UR_{t-2} + a_2 UR_{t-3} + a_3 UR_{t-4} + \quad (18)$$

$$a_4 LR_{t-2} + a_5 LR_{t-3} + a_6 LR_{t-4} + a_7 ExpVol_{t-2} + a_8 ExpVol_{t-3} + a_9 ExpVol_{t-4} +$$

$$a_{10} Conf_{t-2} + a_{11} Conf_{t-3} + a_{12} Conf_{t-4} + u_{t-1}$$

where UR is the seasonally adjusted unemployment rate, LR is the long term interest rate, $ExpVol$ is Expenditure volatility index and $Conf$ is the confidence index. This system of equations uses the same independent, dependent and instrumental variables as in Carroll et al. (2011), and is used to replicate their original results and extend the analysis to include more current data. This specification is identical to equation 2 and to equation 7 when $a_2 = a_3 = a_4 = a_5 = 0$.

5.2 Data

In order to pursue the most timely and relevant data, I retrieve Canadian data from authoritative sources. These sources are equivalent to those used by Carroll et al. (2011). Consumption is measured by Final

Consumption Expenditure on Non-Durable Goods and Services in 2012 Chained Dollars from Statistics Canada (Table 36-10-0107-01). The long term interest rate is measured by the average rate of Government of Canada marketable bonds over ten years pulled from Statistics Canada (Table 10-10-0122-01). The unemployment rate is taken from Statistics Canada (Table 14-10-0287-01). As unemployment is available monthly, quarterly unemployment estimates are obtained by taking the average for each quarter. The index of consumption expenditure volatility is calculated as the rolling coefficient of variation over four quarters of the index of Household Final Consumption Expenditure from Statistics Canada (Table 36-10-0106-01), following the formula in the footnote of page 1139 from Carroll et al. (2011).

A standard source for an index of consumer confidence in Canada has been the Conference Board of Canada's consumer confidence index. Following Carroll et al. (2011), Galli (2019) and Kichian and Mihic (2018), I include a consumer confidence index in the regression analysis.

I use two series for consumer confidence. The first index of consumer confidence is by the OECD. The indicator is sourced from the Conference Board of Canada and is harmonized to be comparable between other countries' confidence indices. It is amplitude adjusted to have a long-term average of 100, where deviations above and below the average indicate upswings and downswings of consumer confidence. Because the confidence index is already adjusted to have a fixed mean throughout the series, the latest indicator for each quarter is chosen to capture variation that occurs throughout the months. The second confidence indicator is the consumer confidence index from the Conference Board of Canada. For data before 2002Q1, the consumer confidence index is taken directly from the data provided by Carroll et al. (2011).² For data after 2002Q1, the consumer confidence indicator is taken directly from the Conference Board of Canada website. Each series was rebased such that their base years are 2002Q1 and were combined to give a consumer confidence indicator that spans the entire research time period. This indicator is referred to in this text as the harmonized confidence indicator.

I also include the OECD's Composite Leading Indicator (CLI) as an additional instrument. The CLI is constructed by aggregating a set of component series. For Canada, the component series are: The Deflated M1 monetary aggregate, Industrial Confidence indicator, spread of interest rates, ratio of inventories to shipments and a stock composite index (OECD, 2020c). I employ it in place of the consumer confidence indicator in one specification of the empirical model since it is likely closely related to the consumer confidence

²The Conference Board of Canada only provides data for their consumer confidence index as early as 2002Q1. Carroll et al. (2011) provide their entire dataset online, and therefore I am able to fuse the two datasets together.

index.

5.3 Main Results

Table 2 contains TSLS estimates of the stickiness parameter for various date ranges (using the ivpack software package written by Jiang and Small (2014))³. The first column represents the closest time interval to Carroll et al. (2011) that is available in the data with the preferred specification. The remaining columns are estimates using other date ranges. Table 3 displays the F and effective F statistics (as defined in equation 14) for each of the date ranges. Because the null hypothesis of a weak instrument set cannot be rejected, the confidence intervals are calculated by the CLR test (using the ivmodel software package written by Kang et al. (2020)), which is robust to weak instruments.

Table 2: Stickiness Estimates, Seasonally Adjusted Unemployment Rates

	<i>Dependent variable:</i>			
	ΔC			
	1980 Q1 - 2002Q3	1980Q1-2020Q1	1980Q1-2008Q4	2009Q3-2020Q1
	(1)	(2)	(3)	(4)
χ	0.747*** (0.520, 1.755)	0.790*** (0.578, 1.484)	0.799*** (0.575, 1.561)	0.956** (0.006, 7.695)
Observations	87	157	112	43
Residual Std. Error	0.695 (df = 85)	0.587 (df = 155)	0.661 (df = 110)	0.364 (df = 41)

Note:

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

Carroll et al. (2011) estimate the stickiness parameter to be approximately 0.72, and statistically significant at the 95% confidence level according to the CLR test. This replication is reported in table 2, and is very similar to the results in the original study⁴.

Table 3: F and Effective Statistics for First Stage TSLS regressions for Table 2

Statistic	F	F_{EFF}	Crit. F_{EFF}
1980Q1-2002Q3	3.331	3.699	14.567
1980Q1-2020Q1	4.71	4.689	14.514
2009Q3-2020Q1	1.44	1.295	15.500
Critical Value	11.52		

³All regression tables were formatted using the stargazer package from Hlavac (2018).

⁴Note that the replication only had available data from 1980 for the OECD consumer confidence index. Because of this, the replication can only go as far back as 1980Q1. To keep the estimates comparable across our other subsamples, the sample begins at 1980Q1 throughout the paper. The data have also undergone adjustments from the statistical agencies since the

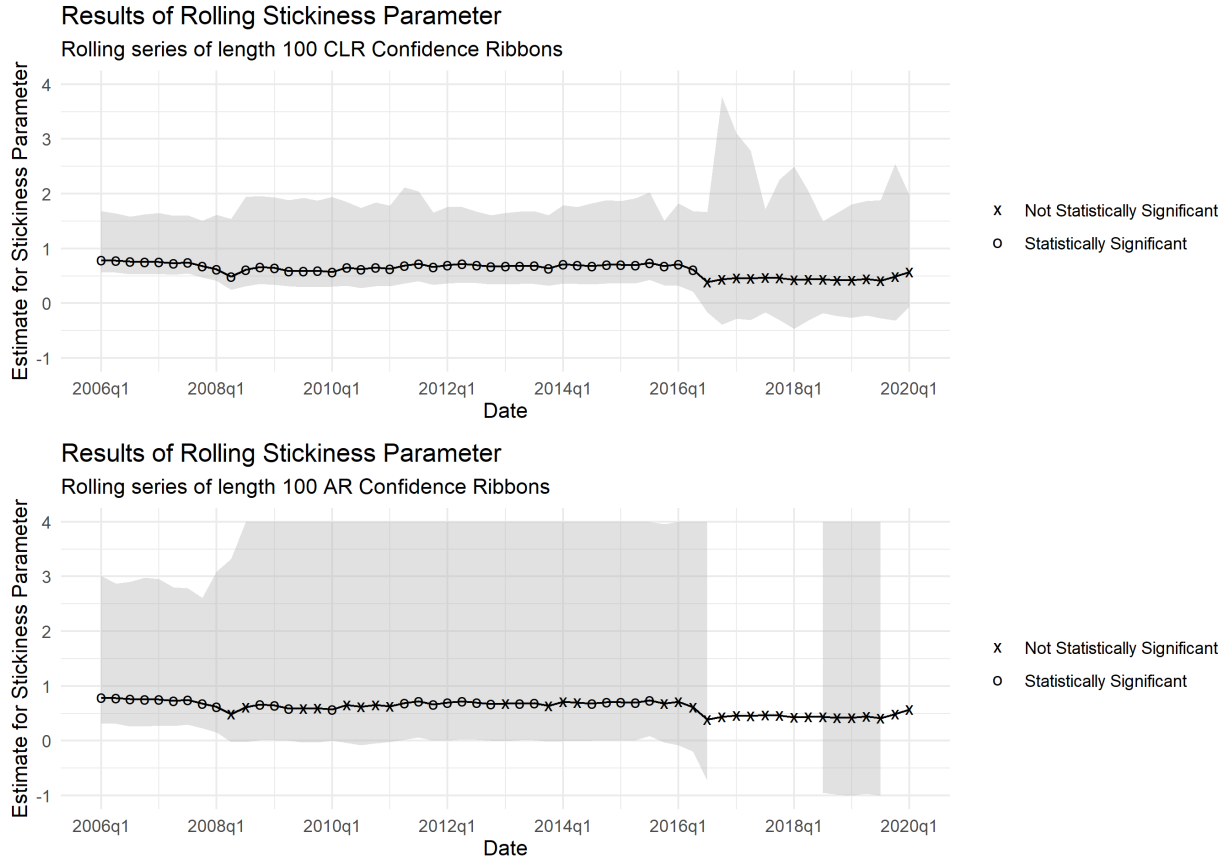
I find a value of 0.75 which is statistically significant at the 95% confidence level. Columns 2 and 3 report estimates with updated subsamples. Column 2 reports the entire sample from 1980Q1-2020Q1, where the stickiness parameter is estimated to be 0.790, suggesting that the stickiness in consumption has increased since Carroll et al.'s initial study. Column 3 contains estimates for the subsample 1980Q1-2008Q4 in order to estimate the stickiness parameter excluding the financial crisis. The estimate is 0.799, higher than the estimate in column 2. The confidence intervals of the estimates in columns 2 and 3 are also tighter than in column 1, which provides further evidence that the stickiness parameter estimates are more precise as the observation period increases. However, it is not clear whether our point estimates for the parameter from columns 1-4 are evidence of a stickiness parameter that is increasing over time, or if they are more precisely estimated as more observations are included. Columns 1-3 are all well within the range of the original stickiness estimate for Canada from Carroll et al. (2011). Column 4 contains estimates for the stickiness parameter following the financial crisis from 2009Q3-2020Q1, and reports a stickiness parameter of 0.956, however the estimate is very imprecise. The CLR interval estimate of column 4 ranges from 0.006 – 7.695, which is economically uninformative, since the stickiness parameter should be bounded between 0 and 1.

Within the broader context of the habit formation empirical literature, Havranek et al. (2017) find that the mean value of habit formation estimates in studies using macroeconomic data is approximately 0.6. For further contextualization, our parameter estimates fall very near to the range to account for popular economic problems such as the equity premium puzzle (0.8) or the hump-shaped response from Fuhrer (2000) (0.8-0.9) (Havranek et al., 2017). An important factor to note is that the CLR confidence interval for each subsample contains values that are greater than one. Economically, the χ coefficient must be at least zero and at most one. Given the presence of a weak instrument set and a relatively small sample size, such large confidence intervals are not unexpected but troubling. Encouraging, however is the lower bound of the estimates in columns 1-3 of table 2. With the lowest lower bound in the estimates being 0.520, the lower bound for stickiness in consumption is economically significant and near the mean estimate for macroeconomic data as reported by Havranek et al. (2017).

I also estimate the stickiness parameters using a rolling subsample with a fixed window. With each estimate, the rolling window moves forward one quarter and removes the earliest quarter from its sample. In order to capture the effect of the financial crisis while keeping a large enough sample size to obtain more

original study from Carroll et al. (2011). These slight adjustments to consumption and different instruments may also have increased the disparity between the replication in this study and Carroll et al.'s original finding. However the difference in parameter estimates are only slight, and nearly all parameter estimates are within a close range of Carroll et al. (2011).

Figure 1: Rolling Parameter Estimates, sample size 100



Note: Confidence sets are left blank when they fall fully in the negative range. This is done to make the graph easier to read.

precise estimates, I begin with a rolling window of 100 observations. I later change the window size to conduct robustness checks. The main result of the rolling estimates is shown in figure 1. Both the CLR and AR estimates are reported in this rolling analysis, in order to provide a more robust interpretation of any changes in the rolling samples over time. See sections 3 and 6.2 for discussions on the differences between the CLR and AR confidence intervals. There are a couple of sharp dips in the rolling estimates, occurring around 2008Q2 (from 0.625 to 0.487), and 2016Q3 (from 0.615 to 0.385). This is consistent with the results reported by Kumar and Jia (2019), who report that stickiness decreases during the financial crisis and then rebounds in a U-shape. However, the confidence intervals around the point estimates are large enough such that it is not possible to definitively say that these point estimates are distinct from one another from these results. Also this effect does not seem robust to changes in the rolling sample size (see section 5.4).

5.4 Robustness Checks

Table 4 contains estimates with seasonally adjusted unemployment rates and the harmonized consumer confidence index. Unsurprisingly, this specification's estimates are the closest to Carroll et al., since their original confidence index data were used from 1980Q1 to 2002Q1 and combined with the new consumer confidence index data from the Conference Board of Canada for the 2002-2020 portion of our sample. It is considerably more volatile than the OECD's amplitude-adjusted version by definition since there are no amplitude adjustments. The estimates from this series decrease as the sample size increases, with the exception of the 2009Q3-2020Q1 sample in column 4, however its upper bound is nearly 3 times higher than the next highest upper bound in the table.

Table 4: Estimates with 'harmonized' confidence index

	<i>Dependent variable:</i>			
	ΔC			
	1980 Q1 - 2002Q3	1980Q1-2020Q1	1980Q1-2008Q4	2009Q3-2020Q1
	(1)	(2)	(3)	(4)
χ	0.726** (0.194, 2.542)	0.688** (0.168, 2.421)	0.690** (0.173, 2.026)	1.637** (0.591, 7.405)
Observations	91	173	116	43
Residual Std. Error	0.679 (df = 89)	0.572 (df = 171)	0.624 (df = 114)	0.458 (df = 41)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

Table 5: Revised Estimates, Original Sourced Data, First Differenced Instruments

	<i>Dependent variable:</i>			
	ΔC			
	1980 Q1 - 2002Q3	1980Q1-2020Q1	1980Q1-2008Q4	2009Q3-2020Q1
	(1)	(2)	(3)	(4)
χ	0.690*** (0.309, 1.973)	0.726*** (0.366, 1.572)	0.810*** (0.443, 1.923)	0.388 (-Inf, Inf)
Observations	86	156	111	43
Residual Std. Error	0.676 (df = 84)	0.570 (df = 154)	0.663 (df = 109)	0.334 (df = 41)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01			

This is also evident in table 5, where the first difference of each instrument is used. This specification is robust against differences in index base years in the instruments by looking at the changes in unemployment

rates, expenditure volatility, long-term interest rates and consumer confidence. The magnitudes of the stickiness parameter decrease when the instruments are first differenced, however all but the post-financial-crisis samples are statistically significant at the 95% confidence level. The confidence intervals span wider than in the original specification, and the post-financial crisis sample in column 4 has an unbounded confidence interval. As is explained in Dufour, Khalaf, and Kichian (2006) and in Dufour (2003), wider confidence sets can be seen as evidence of poor model fit or problems with model identification. This suggests that the levels of instruments are likely more informative than first-differenced instrument sets. The unbounded confidence interval in column 4 of table 5 implies identification problems with this specific instrument set (Dufour, 2003).

Table 6 contains estimates where consumer confidence is replaced with the OECD's Composite Leading Indicator. The results in table 6 are closest to the preferred specification in table 2, and the confidence intervals are the tightest among the robustness check specifications, with the exception of the estimate in column 4 in table 6, which contains a large confidence interval and a point estimate outside of economical values. The imprecise estimates in column 4 are present in all of the specifications, and are good evidence of the importance of having enough data to estimate consumption stickiness.

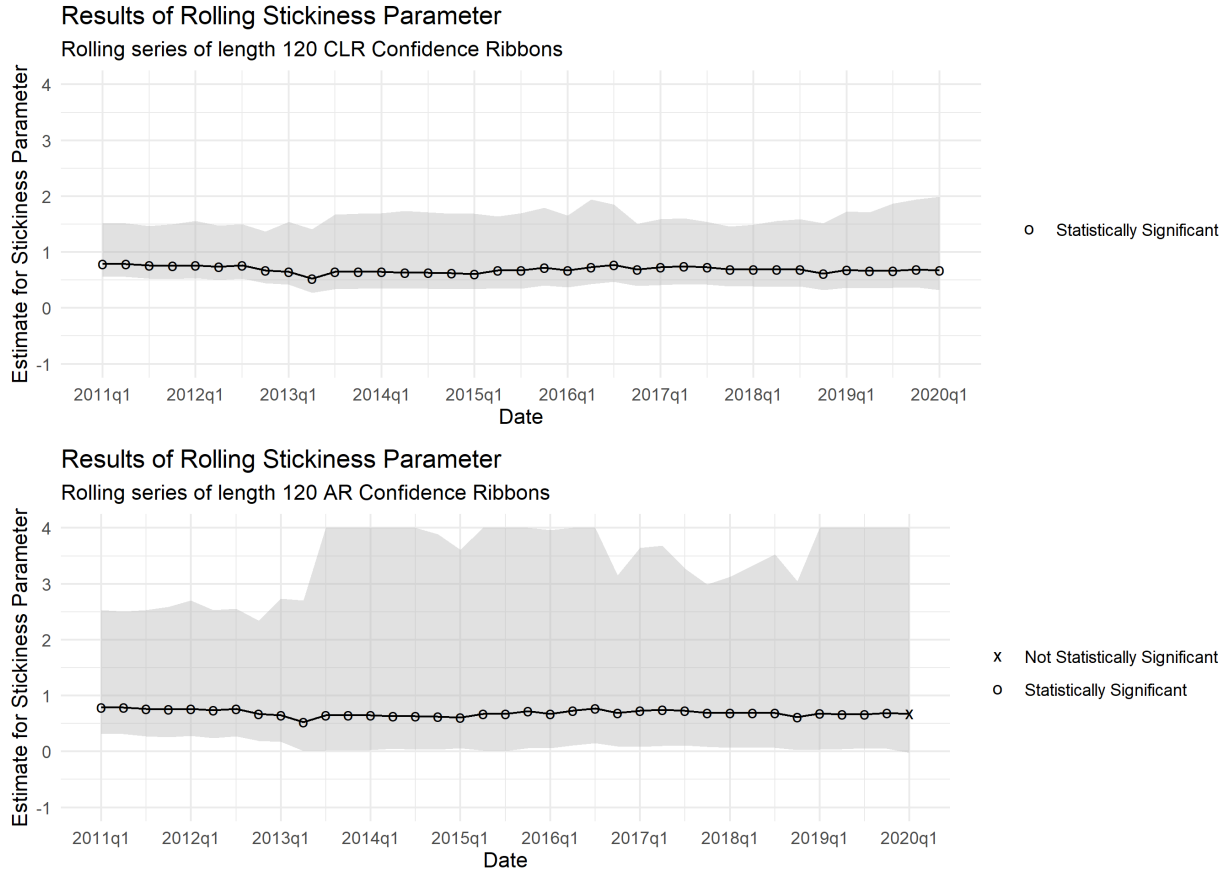
Table 6: Composite Leading Indicator from OECD

	<i>Dependent variable:</i>			
	ΔC			
	1980 Q1 - 2002Q3	1980Q1-2020Q1	1980Q1-2008Q4	2009Q3-2020Q1
	(1)	(2)	(3)	(4)
χ	0.697*** (0.426, 1.597)	0.763*** (0.517, 1.417)	0.747*** (0.492, 1.440)	1.301*** (0.382, 9.139)
Observations	91	173	116	43
Residual Std. Error	0.670 (df = 89)	0.591 (df = 171)	0.639 (df = 114)	0.405 (df = 41)

Note: *p<0.1; **p<0.05; ***p<0.01

In order to verify the robustness of the rolling estimates obtained in figure 1, I repeat the rolling estimates with larger rolling sample window of 120 and a smaller rolling sample of 87, which are shown in figures 2 and 3 respectively. The robustness checks in figures 2 and 3 cast some doubt on the results from figure 1. In figure 3, the smaller sample size negatively impacts the precision of the estimates, especially according to the AR confidence intervals test. Of note is the fact that the dip in the stickiness parameter around 2008 is no longer evident, and both the AR and CLR confidence sets find the estimates around 2008 to be

Figure 2: Rolling Parameter Estimates, sample size 120.



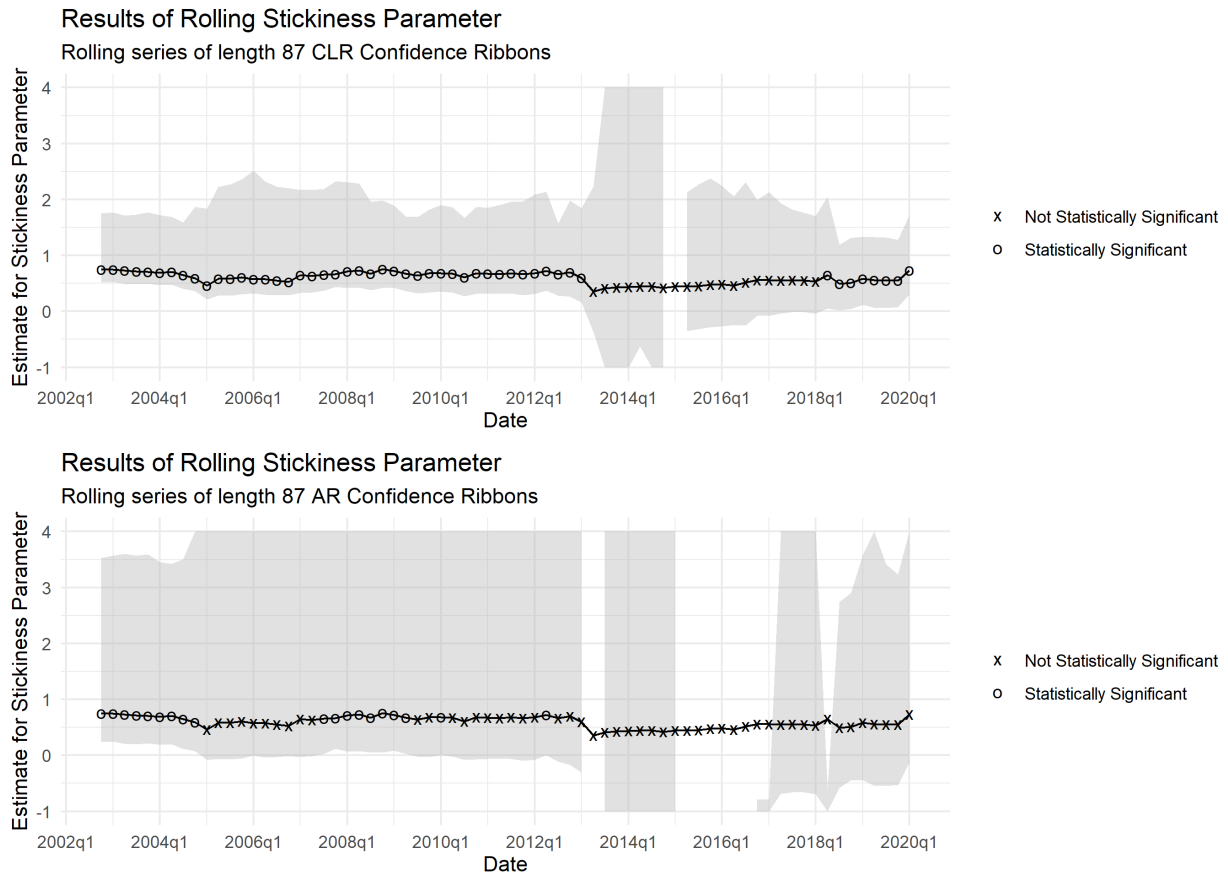
Note: Confidence sets are left blank when they fall fully in the negative range. This is done to make the graph easier to read.

statistically significant. The sample size in figure 2 is large enough that the effect of the financial crisis is already enveloped in the first estimate, however statistical significance is held throughout the entire rolling estimate set.

5.5 Stationarity of Main Specification

The estimates in both the main results and the robustness checks all contain upper bounds in the confidence interval that are greater than one. If it were the case that the stickiness parameter is in fact greater than or equal to one, then the series of changes in the log of consumption would have a unit root and therefore be non-stationary. To test for this, I employ the Augmented Dickey-Fuller test to test for stationarity of the variable $\ln\Delta C_t$, the left-hand side variable of equation 17 (Dickey & Fuller, 1979). I include 4 lags of the dependent variable in order to control for potential autocorrelation. The Augmented Dickey Fuller

Figure 3: Rolling Parameter Estimates, sample size 87.



Note: Confidence sets are left blank when they fall fully in the negative range. This is done to make the graph easier to read.

test statistic for the main specification is -4.273, which rejects the null hypothesis of a unit root at the 1% confidence level. This is strong evidence that the dependent variable in equation 17 is stationary, implying that it is unlikely for the stickiness parameter to be greater than or equal to one. Therefore the large confidence intervals are likely a symptom of weak identification of the model rather than evidence that the series is non-stationary.

6 Discussion

6.1 Macroeconomic Considerations

The parameter χ is the stickiness parameter estimated in section 5 and has the same interpretation as in equation 3. In the habit formation model, point estimates of χ that are greater than 0.5 imply that consumers derive most of their utility in reference to past consumption. Using the estimate in column 2 of table 17 (0.79) as χ , the utility function in equation 3 appears as $U(0.21C_t + 0.79\Delta C_t)$, suggesting that utility is strongly non-time-separable.

The main results from the estimates above are consistent with the average results in the literature (Havranek et al., 2017). These estimates suggest that consumption is likely sticky and are therefore consistent with excess smoothness and hump-shaped responses as found in Fuhrer (2000). The point estimates are relatively high. Under the habit formation model, this implies that consumers generate a significant amount of utility from the change in consumption from the previous period rather than the level. Under the sticky expectations model, this implies that consumers are on aggregate largely inattentive and infrequently update their expectations around economic conditions.

Having an accurate idea of the stickiness parameter is crucial for policymakers. Fuhrer (2000) uses vector autoregression models to show that a proper calibration of habit formation improves a model's description of consumption and inflation dynamics relative to reality. Amato and Laubach (2004, p. 322) examine the impact of consumption stickiness when a policymaker implements an interest rate rule of the form $i_t = ai_{t-1} + b\pi_t + cy_t$, where i_t is the nominal interest rate, π_t is inflation measured by the growth rate of the aggregate price index and y_t is deviation from aggregate output from its steady state level. Through a general equilibrium model with different calibrations for habit formation, Amato and Laubach (2004) find that the weights on the nominal interest rate and inflation, the parameters a and b respectively,

are decreasing as the stickiness parameter increases, with the weight on inflation dropping significantly as stickiness increases. While an analysis of marginal changes in the habit formation parameter is not present, these studies characterize the importance of having an accurate estimate of consumption stickiness, as understanding the dynamics of consumption and spending can help optimally shape policy based on the aggregate consumption response.

A key point to consider in these estimations is the time span and changing economic conditions over the sample. Two of the instruments used in the estimation, the long-term interest rate and volatility of final household consumption expenditure index, change dramatically between 1980 and 2020. The variation in both instruments was much higher prior to 1990. As is evident in figure 4, the long term interest rate has been steadily declining since 1990. The declining interest rates is especially interesting, since the simplified model above assumes that the interest rate for a consumer is constant, consistent with the framework by Dynan (2000). Through figure 4, it is clear that there are several long-term patterns in the long term interest rates. There is a period of approximately 15 years where the rates increase steadily, and a period of approximately 30 years where they decline steadily. It is possible that the changes in economic conditions affected the quality of the fit of some specifications. The large confidence intervals after 2016 in figure 1 could potentially reflect a low quality of fit from the instruments.

Note that the main specification assumes constant interest rates when this assumption does not hold in reality. A rational consumer who foresees steadily declining interest rates, and therefore a decreased value in saving, may behave very differently than a consumer that foresees climbing interest rates, and both would behave differently than a consumer expecting constant interest rates. A falling interest rate over time may induce a consumer to steadily increase consumption and decrease savings. A consumer in a habit formation model however, would gradually respond to income shocks, and exhibit smoothing of consumption over time (Dynan, 2000; Fuhrer, 2000). A model that assumes habit formation and a constant interest rate may not be able to distinguish between steadily changing interest rates and habit formation when it assumes that interests rates are constant.

It is important to note that there is no reason that the stickiness parameter should remain static over time. Economic conditions could cause aggregate consumer behaviour to change over time, which could be reflected in the consumption stickiness estimates. Luo et al. (2017) show in their rational inattention model that when attention is elastic, consumers will devote more energy to attention when income is uncertain. To show this, the authors adjust their model to allow for variations in the information processing capacity with

a fixed marginal cost instead of fixing capacity as an exogenous parameter. They find that during periods where there are large economic shocks, it is reasonable to suppose that consumers pay more attention due to the increased volatility in financial conditions. This factor could also change dynamically through the time series analyzed here, since volatility of the expenditure index and interest rates have visibly decreased from 1980-2020. As is remarked upon in Luo et al. (2017), degrees of informational rigidities are not necessarily structural parameters, but can vary depending on the macroeconomic conditions.

This intuition has a parallel in Carroll et al.'s (2020) sticky expectations model. Decreasing volatility in interest rates and expenditures would, in theory, decrease the need and impact for consumers to update their information and expectations. With the exception of the financial crisis in 2008-2009 and more recently, the COVID-19 pandemic in 2020Q1, stable conditions would reduce the distinguishability between updaters and non-updaters, since low volatility in financial conditions would result in expectations between periods to be more similar. A perceived increase in non-updaters would appear as a higher stickiness parameter in the absence of variable interest rates and economic conditions. On the other hand, large sudden shocks such as the financial crisis in 2008 or the oil price decline in 2014-2016, may increase the consumers' attention to macroeconomic conditions. From equation 4, the high profile of large economic shocks may increase the proportion of consumers who update information, and could induce a change in the stickiness parameter. This suggests that a consumer's decision to update information may be endogenous; something that Luo et al. (2017) can reflect in their model.

These models of rational inattention and sticky expectations would theoretically predict that high-profile shocks to economic conditions can theoretically cause consumers' behaviour to shift. This is a possible theoretical justification for the dips in the stickiness parameter during the Financial Crisis found by Kumar and Jia (2019). However, the rolling estimates of consumption stickiness in this study do not provide empirical evidence for these phenomena.

Carroll et al. (2020) formulate a model of sticky expectations that, when tested on simulated microeconomic and macroeconomic datasets generated by their models, seems to match the disparities in estimations reported by Havranek et al. (2017). A key assumption by Carroll et al. (2020) in their simulations is that aggregate income shocks matter less than idiosyncratic shocks by two orders of magnitude. Given the impact of large aggregate shocks on the economy have on stickiness parameters, it may be worthwhile to examine how microeconomic sticky expectations would change in the wake of very large economic shocks.

Another potential cause that could lead to a shifting stickiness parameter is unaccounted shifting demographic parameters (Attanasio & Weber, 1993). Over the course of forty years, shifts in taste parameters as well as the general basket of goods consumed by the typical consumer may have changed drastically. Attanasio and Weber (1993) conduct analysis of the relationship between consumption growth and the real interest rate and find that results are not robust to differences in estimates for differences in cohorts. Additionally, it is possible that tastes have shifted dramatically from the first twenty years of the sample to the second twenty years. If the consumers in the second half of the sample are purchasing entirely different products that have different product characteristics (i.e. longevity, required updates, etc.), then consumers' purchasing habits may change. If the differences in point estimates over time are not solely due to better identification of the models as the sample size increases, this phenomenon may partially explain the change in stickiness parameters. However, this theory is not likely able to explain sudden jumps in parameter estimates.

The dip into statistical non-significance during 2016Q3 in figure 1 and during 2014 in figure 3 is more puzzling. Such a dip in consumer confidence was not found by Kumar and Jia, suggesting that this impact may have been more intense for Canada. One possible candidate for such a dip is the fall in oil prices from 2014-2016. Oil is a regionally intense source of income for provinces like Alberta and Saskatchewan, and the steady fall in oil prices from the end of 2014 and reaching a trough in 2016Q1 may be able to explain such a dip in the stickiness parameter.

6.2 Econometric Considerations

The significance at the 95% confidence level in the rolling estimations in figure 1 represent some interesting phenomena. The CLR confidence interval for each of the rolling parameter estimates implies statistical significance at the 95% level, however the AR confidence intervals report some estimates to be not statistically significant or nearly so. This reflects the different assumptions between the CLR and AR tests as mentioned in section 3. While the AR test is inefficient for our overidentified model, it is robust to excluded instruments, unlike the CLR test. Given the complexity of aggregate economic behaviour and the very large temporal window in this analysis, it is practically impossible to be sure that no instruments are excluded. For this reason, I view the CLR confidence intervals to be the best case scenario for inference with weak instruments, and the AR confidence intervals to be the conservative estimates. Following Dufour et al. (2006), very wide AR confidence intervals can be used to interpret the model fit, since unbounded sets occur when there are difficulties in model identification. From figure 1, increasing confidence bands in the AR case imply

Figure 4: Yield of Long-term bonds, quarterly, Canada



Source: Statistics Canada, Table 10-10-0122-01

a breakdown of model fit, which occur during and after the financial crisis. A possible explanation is the inclusion of the financial crisis in the sample set for all estimates after 2008. The AR confidence bands inflate after the financial crisis, and hover just around 0 for most estimates after 2008. This implies that the model with the current instrument set is no longer well identified. In the future, a formal analysis of identification of sticky consumption models over time would be most useful, since proper comparisons of stickiness parameter estimates in different time periods would be helped by similar levels of identification in the models.

An important distinction to make between the specifications is the existence of smoothing in the OECD confidence index and the unemployment rate when seasonally adjusted. As is evident in equation 10, the smoothness of the data plays an important role when instrumenting for lagged consumption, since a key characteristic of stickiness in consumption is the smooth response to shocks in income. As is evident in the different estimates between table 2 and 4, the presence of smoothed instruments versus their volatile equivalents can have a significant impact on the point estimates.

As Havranek et al. (2017) note in their meta-analysis, the frequency of the data used to estimate consumption stickiness matters. Since smoothing consumption growth is by necessity a temporal concept,

Figure 5: Monthly Average Oil Spot Prices

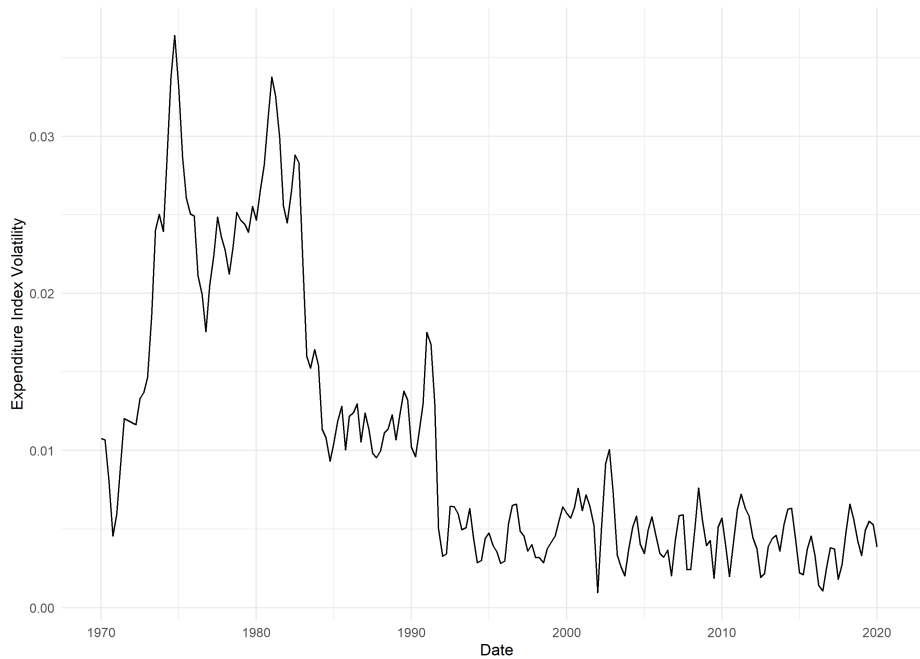


Source: US Energy Information Administration

the frequency of the data warrants significant attention. This is especially important given that much of the data provided by statistical agencies are now provided monthly. To convert the data into a quarterly figure, some aggregation metric must be used, and the results from this study implies that aggregation that leads to smoothing the monthly data may cause the estimates to change.

As is evident from table 2, the point estimates become more precise as more data are used in the estimation. This is not necessarily surprising, however the estimates in column 4, including nearly 11 years of quarterly data, are extremely imprecise. There is a clear need for longer time horizons than 10 years for estimating consumption stickiness, however longer time horizons increase the likelihood that cultural consumption patterns may change (as discussed in section 6.1), or that there may be a structural break in the time series. That the results from the rolling estimates often show different changes in consumption stickiness with different sample sizes is more evidence of the trade-off of including more data but also introducing possible changes in the underlying dataset. This is a factor that should be considered in future analysis of consumption stickiness.

Figure 6: Index of Expenditure Volatility



Source: Statistics Canada, Table 36-10-0106-01. Calculation formula from Carroll et al. (2011)

7 Conclusion

In this Major Research Paper, I update the estimates of consumption stickiness for Canada using aggregate quarterly data from 1980Q1 to 2020Q1, following the specifications from Carroll et al. (2011). The results are consistent with those of Carroll et al. and are generally statistically significant, however the confidence intervals from the weak-instrument-robust estimates imply a wide range of possible values, including some that are beyond reasonable economic values. When conducting a rolling estimation of consumption stickiness, I find that consumption stickiness possibly follows a U-shaped pattern, dipping during the financial crisis of 2008-2009 and returning to previous levels afterward. This is consistent with Kumar and Jia (2019). After the financial crisis, I find evidence of more serious identification issues with the current instrument set. These results shed light on the possible imprecision of the point estimates reported in the literature and show the value in reporting identification-robust confidence intervals.

The results in this Major Research Paper provide avenues for exploration in the literature. Investigation of how salience of economic information may induce changes in the proportion of consumers who are attentive to information (as modelled in Carroll et al. (2020)) may provide insight on the microfoundations of sticky consumption. Second, the assumption of constant interest rates employed in the approximation

of consumption stickiness as an AR(1) equation clearly does not match up to the data. A fruitful area of research may be to analyze intertemporal substitution of consumption when allowing consumption to be sticky. While Everaert and Pozzi (2014) incorporate interest rates in their analysis, it may be worthwhile to conduct this analysis with quarterly data. The importance of a sufficient sample size and strong instruments are highlighted in this analysis, and more careful consideration of model identification over long time horizons should be further investigated.

References

- Amato, J. D., & Laubach, T. (2004). Implications of habit formation for optimal monetary policy. *Journal of Monetary Economics*, 51(2), 305–325.
- Anderson, T. W., & Rubin, H. (1949, 03). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics*, 20(1), 46–63.
- Andrews, D. W., Moreira, M. J., & Stock, J. H. (2006). Optimal two-sided invariant similar tests for instrumental variables regression. *Econometrica*, 74(3), 715–752.
- Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11(1), 727–753.
- Attanasio, O. P., & Weber, G. (1993). Consumption growth, the interest rate and aggregation. *The Review of Economic Studies*, 60(3), 631–649.
- Campbell, J., & Deaton, A. (1989). Why is consumption so smooth? *The Review of Economic Studies*, 56(3), 357–373.
- Campbell, J., & Mankiw, N. G. (1989). Consumption, income, and interest rates: Reinterpreting the time series evidence. *NBER macroeconomics annual*, 4, 185–216.
- Carroll, C. D., Crawley, E., Slacalek, J., Tokuoka, K., & White, M. N. (2020). Sticky expectations and consumption dynamics. *American economic journal: macroeconomics*, 12(3), 40–76.
- Carroll, C. D., Slacalek, J., & Sommer, M. (2011). International evidence on sticky consumption growth. *Review of Economics and Statistics*, 93(4), 1135–1145.
- Conference Board of Canada. (2020). *Index of consumer confidence, all respondents, canada*.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427–431.
- Dufour, J.-M. (2003). Identification, weak instruments, and statistical inference in econometrics. *Canadian Journal of Economics/Revue canadienne d'économique*, 36(4), 767–808.
- Dufour, J.-M., Khalaf, L., & Kichian, M. (2006). Inflation dynamics and the new keynesian phillips curve: An identification robust econometric analysis. *Journal of Economic Dynamics and Control*, 30(9), 1707 - 1727. (Computing in economics and finance)
- Dufour, J.-M., & Taamouti, M. (2005). Projection-based statistical inference in linear structural models with possibly weak instruments. *Econometrica*, 73(4), 1351–1365.
- Dynan, K. E. (2000). Habit formation in consumer preferences: Evidence from panel data. *American Economic Review*, 90(3), 391–406.
- Everaert, G., & Pozzi, L. (2014). The predictability of aggregate consumption growth in oecd countries: A

- panel data analysis. *Journal of Applied Econometrics*, 29(3), 431-453.
- Everaert, G., Pozzi, L., & Schoonackers, R. (2017). On the stability of the excess sensitivity of aggregate consumption growth in the usa. *Journal of Applied Econometrics*, 32(4), 819-840.
- Fuhrer, J. C. (2000). Habit formation in consumption and its implications for monetary-policy models. *American Economic Review*, 90(3), 367-390.
- Fuhrer, J. C. (2017). Expectations as a source of macroeconomic persistence: Evidence from survey expectations in a dynamic macro model. *Journal of Monetary Economics*, 86(C), 22-35.
- Galli, A. (2019). Sticky consumption and wealth effects in switzerland. *Review of Income and Wealth*, 65(4), 930-952.
- Greene, W. H. (2018). *Econometric analysis*. Pearson Education.
- Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of political economy*, 86(6), 971-987.
- Havranek, T., Rusnak, M., & Sokolova, A. (2017). Habit formation in consumption: A meta-analysis. *European Economic Review*, 95, 142-167.
- Hlavac, M. (2018). stargazer: Well-formatted regression and summary statistics tables [Computer software manual]. Bratislava, Slovakia. Retrieved from <https://CRAN.R-project.org/package=stargazer> (R package version 5.2.2)
- Jiang, Y., & Small, D. (2014). ivpack: Instrumental variable estimation. [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=ivpack> (R package version 1.2)
- Kang, H., Jiang, Y., Zhao, Q., & Small, D. (2020). ivmodel: Statistical inference and sensitivity analysis for instrumental variables model [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=ivmodel> (R package version 1.8.1)
- Kichian, M., & Mihic, M. (2018). How important are wealth effects on consumption in canada? *Canadian Journal of Economics/Revue canadienne d'économique*, 51(3), 784-798.
- Kumar, S., & Jia, P. (2019). Financial crisis and persistence: evidence from sticky expectations consumption growth model. *Applied Economics*, 51(17), 1799-1807. Retrieved from <https://doi.org/10.1080/00036846.2018.1529394> doi: 10.1080/00036846.2018.1529394
- Luo, Y. (2008). Consumption dynamics under information processing constraints. *Review of Economic Dynamics*, 11(2), 366-385.
- Luo, Y., Nie, J., Wang, G., & Young, E. R. (2017). Rational inattention and the dynamics of consumption and wealth in general equilibrium. *Journal of Economic Theory*, 172, 55 - 87.
- Montiel Olea, J. L., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3), 358-369.

- Moreira, M. J. (2003). A conditional likelihood ratio test for structural models. *Econometrica*, 71(4), 1027–1048.
- Muellbauer, J. (1988). Habits, rationality and myopia in the life cycle consumption function. *Annales d'Économie et de Statistique*(9), 47-70.
- OECD. (2020a). *Composite leading indicator (cli) (indicator)*. doi: <https://doi.org/10.1787/4a174487-en>
- OECD. (2020b). *Consumer confidence index (cci) (indicator)*. doi: <https://doi.org/10.1787/46434d78-en>
- OECD. (2020c). *Oecd composite leading indicators:turning points of reference series and component series* (Tech. Rep.). Author. Retrieved from <http://www.oecd.org/sdd/leading-indicators/CLI-components-and-turning-points.pdf>
- Pagel, M. (2017). Expectations-based reference-dependent life-cycle consumption. *The Review of Economic Studies*, 84(2), 885–934.
- Pflueger, C. E., & Wang, S. (2015). A robust test for weak instruments in stata. *The Stata Journal*, 15(1), 216–225.
- Sommer, M. (2007). Habit formation and aggregate consumption dynamics. *The BE journal of macroeconomics*, 7(1).
- Statistics Canada. (2020a). *Table 10-10-0122-01 financial market statistics, last wednesday unless otherwise stated, bank of canada*.
- Statistics Canada. (2020b). *Table 14-10-0287-0 labour force characteristics by province, monthly, seasonally adjusted*.
- Statistics Canada. (2020c). *Table 36-10-0106-01 gross domestic product price indexes, quarterly*.
- Statistics Canada. (2020d). *Table 36-10-0107-01 household final consumption expenditure, quarterly, canada (x 1,000,000)*.
- Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4), 518–529.
- U.S. Energy Information Administration. (2020). *Spot prices, crude oil in dollars per barrel, products in dollars per gallon*. Retrieved from https://www.eia.gov/dnav/pet/pet_pri_spt_s1_m.htm