Staff Discussion Paper/Document d’analyse du personnel —
2020-13

Last updated: December 1, 2020

Toward a HANK Model for Canada: Estimating a Canadian Income Process

by Iskander Karibzhanov

International Economic Analysis Department
Bank of Canada, Ottawa, Ontario, Canada K1A 0G9

ikaribzhanov@bankofcanada.ca

ISSN 1914-0568

Acknowledgements

The author thanks the workshop participants from the Bank of Canada's International Economic Analysis Department and Markus Poschke of McGill University for helpful comments. Special thanks go to Maren Hansen for editing the note and improving its readability.

Abstract

I study individual earnings dynamics using panel data on Canadian workers. I first show that, similar to US findings, the distribution of Canadian income growth is leptokurtic. To generate such high kurtosis, I use a common continuous-time specification of the individual earnings as a stochastic process with a random (Poisson) arrival of normally distributed jumps. The fitted earnings process matches the eight targeted moments well. The estimated parameter values are consistent with the existence of both transitory and persistent components in earnings. On the methodological side, I show how the estimation process can be accelerated significantly by parallelizing Monte Carlo simulations on graphical processing units with massive savings in computational time. My estimates represent a key first step in developing quantitatively realistic Heterogeneous Agent New Keynesian (HANK) models for the Canadian economy. HANK models are important tools for understanding consumption behaviour and analyzing the transmission mechanism for monetary policy. The estimated process in this paper may prove useful in other contexts where an empirically realistic representation of household earnings dynamics is vital.

Bank topic: Economic models; Labour markets
JEL codes: D31, E24, J31

1. Introduction

Recently, Kaplan, Moll and Violante (2018) introduced a Heterogeneous Agent New Keynesian (HANK) model to explore the quantitative implications of household heterogeneity for the transmission of monetary policy. While keeping the nominal rigidities that characterize standard New Keynesian models, this model replaces the representative agent with heterogeneous households that face uninsurable earnings risk. This heterogeneity gives rise to inequality in income, wealth and consumption, which in turn affect aggregate demand. The authors find that household heterogeneity plays an important role in matching facts about consumption behaviour and may account for many of the stylized facts that do not seem consistent with standard New Keynesian models.

In this paper, I estimate the Canadian earnings process. This is a key first step in the development of quantitatively realistic HANK models of the Canadian economy. These models are important for studying consumption behaviour and the mechanism for the transmission of monetary policy. First, I identify the type of earnings shocks households face in Canada. Specifically, I find that the distribution of earnings shocks is leptokurtic, meaning that it is characterized by large, infrequent shocks rather than small but frequent ones. The frequency and size of earnings shocks are crucial in households' decisions around saving and the degree of liquidity desired in assets. Hence, an empirically realistic representation of household earnings dynamics is vital for calibrating the relative holdings of liquid and illiquid assets in HANK models. This process, in turn, plays a key role in consumption dynamics and the transmission of monetary and other shocks.

On the methodological side, I show how the estimation process can be significantly accelerated, making it possible to run millions of simulations in under a second. This involves parallelizing the code to run on multiple graphical processing units (GPUs), in addition to using one-pass algorithms to utilize GPU resources more efficiently. My codes for implementing these procedures are publicly available and should be helpful to other researchers working with HANK models.

In Section 2, following Guvenen et. al. (2015), I use a high-quality panel earnings dataset from Statistics Canada and find that, as in the United States, in Canada earnings growth displays a very high kurtosis relative to a Gaussian density. My key finding is that Canadian earnings exhibit substantial deviation from lognormality, such as negative skewness and elevated kurtosis. In Section 3, I follow Kaplan, Moll and Violante (2018) by defining the continuous-time stochastic process for the earnings process. In Section 4, I target the eight key moments of the distribution to estimate the earnings process using a simulated method of moments. In Section 5, I expand on the methodological contributions described above. In Section 6, I conclude.

2. Data

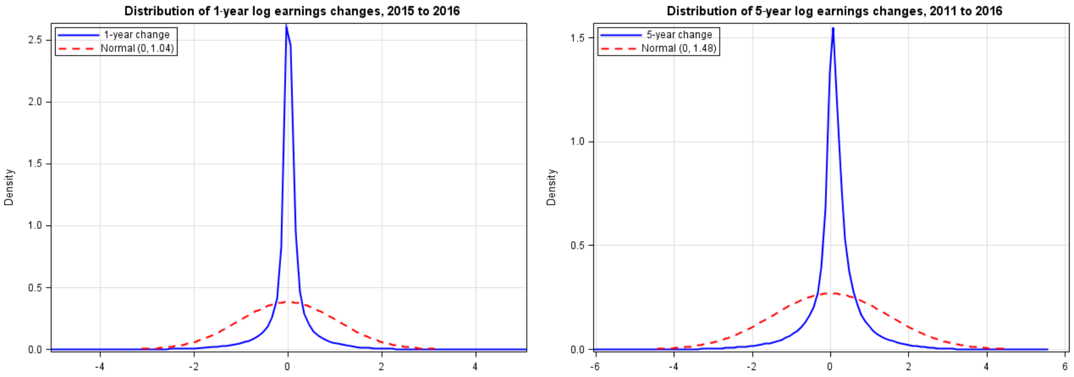
I obtained a custom tabulation from the Longitudinal Administrative Databank (LAD).[1] The LAD is a panel comprising a 20 percent sample of all annual tax filings between 1982 and 2016. As in Guvenen et al. (2015), the sample is restricted to males of working age who have recently participated in the labour market, earning above minimum wage.[2]

The eight key moments of the distribution of earnings growth rates in Canada are close to the moments Guvenen et al. (2015) obtained from the Social Security Administration data for the United States (**Table 1**). We can see in **Chart 1** that the distribution of changes in earnings exhibits substantial deviations from lognormality, such as very high kurtosis. In particular, the empirical density plotted in the left panel of **Chart 1** has a kurtosis of 13, which is much higher than that of a normal distribution, with a kurtosis of 3.

Table 1: Key moments of the earnings process in the United States and Canada

Moment	US data	Canadian data	Model estimates
Variance: annual log earnings	0.70	0.760	0.760
Variance: 1-year changes	0.23	0.217	0.215
Variance: 5-year changes	0.46	0.437	0.439
Kurtosis: 1-year changes	17.8	13.38	13.36
Kurtosis: 5-year changes	11.6	8.782	8.777
Share of 1-year changes < 10%	0.54	0.51	0.524
Share of 1-year changes < 20%	0.71	0.68	0.650
Share of 1-year changes < 50%	0.86	0.85	0.836

Chart 1: The empirical density of one-year and five-year log earnings changes



Source: Statistics Canada, Longitudinal Administrative Databank, 2015 to 2016.

---

[1] The custom tabulation is produced by the Income Statistics Division of Statistics Canada.

[2] "Recently participated" means they must have been earning above the minimum wage in the previous year and during at least two of the four years before that.

3. The stochastic income process

Continuous time provides a natural and parsimonious approach to model an individual earnings process with leptokurtic annual income growth: the random (Poisson) arrival of normally distributed jumps generates kurtosis in data observed at discrete time intervals.

Following Kaplan, Moll and Violante (2018), I decompose quarterly log earnings, $\log z_{it}$, into two independent components, a transitory shock, $z_{1,it}$, and a permanent shock, $z_{2,it}$:

$$\log z_{it} = z_{1,it} + z_{2,it}.$$

The intuition behind this decomposition is that we want to model not only periodic temporary changes in income that occur within an individual's career, but also infrequent, large and persistent career shocks (e.g., switching to a higher-paying job).

Both processes $z_{j,it}$ are constructed as simple jump-drift processes, which sometimes jump but have a continuous deterministic evolution that drifts toward zero at rate δ_j between the jump times. More specifically, both processes are given by:

$$dz_{j,it} = -\delta_j z_{j,it} dt + dJ_{j,it},$$

where $dJ_{j,it}$ captures jumps arriving at a Poisson rate λ_j. Thus, conditional on a jump, a new log earnings state $z'_{j,it}$ is drawn from a normal distribution with mean zero and variance σ_j^2, $z'_{j,it} \sim N(0, \sigma_j^2)$. The key difference between this continuous-time formulation and a discrete-time AR(1) process is that the arrival of jumps is stochastic with an additional frequency parameter, λ_j.

4. Estimation using a simulated method of moments

I estimate the earnings process through a simulated method of moments using the LAD data discussed above. The estimation targets the eight key moments reported in **Table 1**. The two components of the income process are identified separately using the variance and kurtosis of the distribution of one-year and five-year changes in earnings. Since the LAD data are reported annually, I first simulate quarterly earnings from the model and then aggregate them to annual earnings. The objective is to minimize the weighted sum of differences between simulated moments and those of the data.

I obtain a good estimation fit and confirm that both transitory and persistent components exist in earnings (**Table 2**). The transitory jumps ($j = 1$) arrive on average once every four years with a half-life of around three-quarters; these can be interpreted as periodic temporary shocks. The persistent jumps ($j = 2$) arrive on average once every 357 years with a half-life of around 92 years and can be interpreted as large and persistent "career" shocks. The innovations in both components are relatively large and similarly sized.

Table 2: Parameter estimates of the earnings process

	Parameter	Transitory component j = 1	Persistent component j = 2
Arrival rate	λ_j	0.061	0.0007
Mean reversion	δ_j	0.227	0.002
Standard deviation of innovations	σ_j	1.46	1.93

Note: Rates are expressed as quarterly values.

5. Code optimization

I use several innovations to significantly accelerate the estimation process relative to Kaplan, Moll and Violante (2018). I parallelize estimation to run on GPUs using the CUDA—Compute Unified Device Architecture—programming language. My code optimizations make it possible to evaluate the objective function with 1 million simulations in 20 milliseconds on NVIDIA Tesla V100 GPUs. In comparison, Kaplan, Moll and Violante's original Fortran codes require 90 seconds to do the same when running sequentially, or 6 seconds when running in parallel on 36 CPUs using OpenMP. As a result, I can speed up estimation by a factor of 4,500 over conventional single-threaded code and by a factor of 285 over multi-threaded code. A key challenge in doing so is that the original Fortran code requires 34 gigabytes of memory, which exceeds the memory size limit of a GPU. To solve this problem, I implement a parallel online algorithm for computing third and fourth central moments, as described in Pébay et al. (2016). By using their one-pass formulas to recursively update the variance, skewness and kurtosis, I can avoid storing simulated data in global memory on GPUs, where the per-processor memory bandwidth is extremely small and communication costs are high. My code is available at GitHub and should be of use to other researchers working with HANK models.

6. Conclusion

The distribution of Canadian earnings growth is close to that found for the United States by Guvenen et al. (2015). I confirm their finding that there is substantial deviation from lognormality, such as negative skewness and elevated kurtosis. I also fit the earnings process using a simulated method of moments to target the eight key moments of the distribution. Finally, I improve the efficiency of Monte Carlo simulations in Kaplan, Moll and Violante (2018) by parallelizing the code to run on GPUs.

The estimates in this paper represent a key first step in the development of quantitatively realistic HANK models for the Canadian economy. These models allow us to better understand consumption behaviour, and they offer another tool for researchers and policy-makers to use in analyzing the monetary policy transmission mechanism. The estimated process in this paper may also prove useful in other contexts where an empirically realistic representation of household earnings dynamics is vital.

References

Guvenen, F., F. Karahan, S. Ozkan and J. Song. 2015. "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?" National Bureau of Economic Research Working Paper No. 20913.

Kaplan, G., B. Moll and G. L. Violante. 2018. "Monetary Policy According to HANK." *American Economic Review* 108 (3): 697–743.

Pébay, P., T. Terriberry, H. Kolla and J. Bennett. 2016. "Numerically Stable, Scalable Formulas for Parallel and Online Computation of Higher-Order Multivariate Central Moments with Arbitrary Weights." *Computational Statistics* 31 (4): 1305–1325.