MULTI-VARIATE EARNING DYNAMICS

CHRISTOS MAKRIDIS* & ALEXIS WONG[1]

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The current stochastic process used to model earnings and decompose them into temporary and permanent shocks is extremely important for studying the structure of earnings processes. This actual modeling of the dynamics of the earning process can be enhanced with a multi-variable effect. The large amounts of data that can now be obtained to model earning processes are very vast and correlated with each other, this means that each time we will want to correlate more variables that we see that affect our stochastic process and somehow incorporate them into our study. For the decomposition of transitory and permanent shock, we therefore need a good scheme and a good one to correlate these objects.

Keywords: Earnings processes, Incomplete markets, Partial insurance, Estimation

* *Faculty of Engineering, Universidad de Los Andes, Mérida, Venezuela*

1 INTRODUCTION

We can use vector autoregression strategies and a multi-variable approach to encompass more complex processes. [1].

A current study carried out by Hryshko and Manovskii (2020), indicates a modern earning process modeling based on the decomposition in transitory and permanent shocks, in fact it is an improved version of the one indicated in the study previously done by Friedman (1957) . In this work by Hryshko and Manovskii, they present what can be interpreted as ways to optimize the equation, seen from the point of view of a pure statistician or a mathematician, in addition, we want it to be understood and in this way be used in production by practicing statesman's, graduates and engineers. Hideously complicated as the masterpieces explained by their creators are defined. We have another document that makes an implementation of this model

2 DEFINITIONS

We have in our first approaches to the literature that, in the first place, indicate rigorous mathematical definitions, the first definition of the decomposition in shocks of the earning process is given by:

$$y_{it} = \alpha_i + p_{it} + t_{it}$$
$$p_{it} = \varphi_p p_{it} - 1 + \xi_{it}(1)$$
$$t_{it} = \theta(L)\epsilon_{it},$$

"...where log-earnings y_{it} of individual i at time t consist of the permanent component, p_{it}, and the transitory component, t_{it}. If φ_p is close to 1, the shocks ξ_{it} are highly persistent (and are truly permanent if φ_p is 1), and if $\theta(L) = 1$ (where $\theta(L)$ is a moving average polynomial in the lag operator L), the shocks ϵ_{it} are completely transitory."

— Daly, Hryshko 2020

2.1 Intuitive Definition

DESCRIPTIVE We have seen a change in the way we interpret the model in the following definition:

$$W_{it} = W_{it}^P + W_{it}^T,$$
$$E(W_{it}^P) = E(W_{it}^T) = E(W_{it}^P + W_{it}^T) = 0$$

"...The Friedman's formulation is far too simplistic."

— Cappellari 2010

2.2 Less Abstract Definition

WAY MORE INTUITIVE In a second approach to the literature of this modeling, we have seen a definition for a clearer public or in this case, something more intuitive, at first it could be seen that it is the same as the first quote but really here we give a more intuitive sketch to proceed with our solution.

A TYPICAL EARNING PROCESS

1. y_{it} is individuals i log-earnings residuals at a given t
2. p_{it} is the permanent component (Random Walk)
3. t_{it} is the transitory component (MA(1), ARMA(1,1), AR(1), or iid;)

2.3 Some Equivalences

Now according to the definitions of the different authors we have a description of what the profit process model is good enough to implement a paradigm shift to the current model, we can proceed to see certain similarities through the mathematical language.

PERMANENT COMPONENT We can assume that the permanent component is a basic Random Walk model given by:

$$x_t = x_{t-1} + w_t, \tag{4.18}$$

"...Alternatively (Dickens 2000, Cappellari and Leonardi 2010): Random walk (RW) model."

$$W_{it}^{P} = r_{iexp(t)} = r_{iexp(t-1)} + \xi_{iexp(t)}$$
$$r_{i0\ i}(0, \sigma_r^2); \xi_{iexp(t)\ ie}(0, \sigma_\xi^2)$$

TRANSITORY COMPONENT t_{it} is the transitory component: MA(1), ARMA(1,1), AR(1), or iid; (Hryshko Manovskii, 2020)

THEREFORE Assume low order ARMA, e.g. ARMA(1,1). (Capellari 2010)

3 IMPROVEMENTS AND DISCUSSIONS

"...When you think about someone's income, it is a function of a persistent component (usually thought of as their productivity and ability) and a transitory component (something contemporaneous that hits them). But, data has problems and so there are better ways to do the decomposition than others."

— Makridis 2021

In fact it is totally true, most of the authors have an accurate opinion, there is a lot of literature on how to measure the processes of earning dynamics.

"There is an extensive literature estimating stochastic processes for individual earnings."

— Manovskii 2020

3.1 What Improvements?

UNBALANCED DATA Allowing for an unbalanced sample.

ERROR Allowing for measurement error in income.

MORE VARIABLES Allowing for the incorporation of consumption and/or labor supply (hours worked) data.

Due to the nature of the data, the number of processes in place, and the number of correlations that exist, academics try to balance this and generalize it to values of degrees of difference. Another improvement given is the lack of existence of a public version of a package to calculate the dynamics of the profit directly by some end user.

LACK OF AUDIENCE This is also really one of the inspirations for this work, to facilitate the study of these models for users and researchers in the area.

4 PROPOSED SOLUTION

After analyzing the model we have very well and the proposed improvements, a solution that is consistent enough to be applied can be designed, this design could allow an unbalanced data panel and at the same time multiple variables.

4.1 MARIMA Approach

We can see that the lag (Theta) model is in fact an ARMA model, taking this into account, it means that our decomposition is done from the regression made by this model.

In the literary description, the decomposition of the temporal shock is given by the residuals of the first-order ARMA process. In fact they tell us that we can take on a low-order ARMA at any time.

In the past works, they try to optimize the first order of integration, that is why the model we use in the project is the ARIMA model.

In a complicated process (for example, one with 13 variables) we need to correlate all of them to test the hypothesis and in turn take the residuals of the model that is used, for this we will use the MARIMA model, which is the same ARIMA only with a functionality to multi-variable processes.

SUMMING UP Ultimately, the original model only passes a process with a single variable.

We will modify the model to process data sets that represent more complex earning processes.

4.2 ARMA model automation

The main idea of this part is to automate the ARMA model selection process, taking into account that we have used a MARIMA model.

We will use a model selection automation function. (**autoMARIMA**).

This function will select the best MARIMA model and will include all the coefficients that we need, we assume that decomposing the MARIMA model will give us:

The AR (1) and MA (1) coefficients that we need to assemble the decomposition of the shocks.

IN RESUME Having a trained an automatic MARIMA model, we can decompose the model and obtain all the necessary parameters, in this case, from the **autoMARIMA** we will obtain the residuals of all the variables present and we will also obtain an ARMA (1,1), conveniently.

5 ACTIONS AND RESULTS

We have implemented a direct version of the algorithm in R. Which we call the original version.

DATASET We use an NLSY dataset. The NLSY collects extensive information on respondents' labor market behavior and educational experiences. The survey also includes data on the youths' family and community backgrounds to help researchers assess the impact of schooling and other environmental factors on these labor market entrants. Data from the NLSY also aid in determining how youths' experiences relate to establishing careers, participating in government programs, and forming families.

Table 1: Datasets

Name	Rows	Columns
NLSY simple	70510	13
None	None	2

Reference to Table 1.

5.1 Basic Model

We first implement the basic model described by Daly, Hrysko and Manovskii. With this baseline approach, we process the NLSY data. Please take a look at Figure 1 on the next page.

Reference one of the subfigures as Figure 1b on the following page.

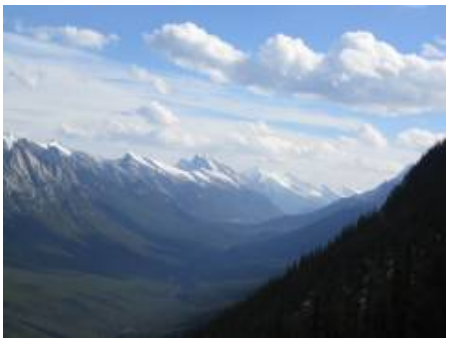
Nulla in ipsum. Praesent eros nulla, congue vitae, euismod ut, commodo a, wisi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Aenean nonummy magna non leo. Sed felis erat, ullamcorper in, dictum non, ultricies ut, lectus. Proin vel arcu a odio lobortis euismod. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Proin ut est. Aliquam odio. Pellentesque massa turpis, cursus eu, euismod nec, tempor congue, nulla. Duis viverra gravida mauris. Cras tincidunt. Curabitur eros ligula, varius ut, pulvinar in, cursus faucibus, augue.

Nulla mattis luctus nulla. Duis commodo velit at leo. Aliquam vulputate magna et leo. Nam vestibulum ullamcorper leo. Vestibulum condimentum rutrum mauris. Donec id mauris. Morbi molestie justo et pede. Vivamus eget turpis sed nisl cursus tempor. Curabitur mollis sapien condimentum nunc. In wisi nisl, malesuada at, dignissim sit amet, lobortis in, odio. Aenean consequat arcu a ante. Pellentesque porta elit sit amet orci. Etiam at turpis nec elit ultricies imperdiet. Nulla facilisi. In hac habitasse platea dictumst. Suspendisse viverra aliquam risus. Nullam pede justo, molestie nonummy, scelerisque eu, facilisis vel, arcu.

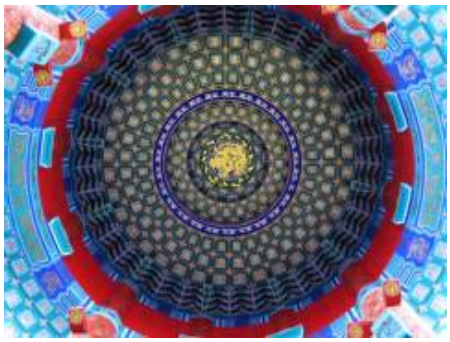
Curabitur tellus magna, porttitor a, commodo a, commodo in, tortor. Donec interdum. Praesent scelerisque. Maecenas posuere sodales odio. Vivamus metus lacus, varius quis, imperdiet quis, rhoncus a, turpis. Etiam ligula arcu, elementum a, venenatis quis, sollicitudin sed, metus. Donec nunc pede, tincidunt in, venenatis vitae, faucibus vel, nibh. Pellentesque wisi. Nullam malesuada. Morbi ut tellus

(a) A city market.

(b) Forest landscape.



(c) Mountain landscape.



(d) A tile decoration.

Figure 1: A number of pictures with no common theme.

ut pede tincidunt porta. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam congue neque id dolor.

Donec et nisl at wisi luctus bibendum. Nam interdum tellus ac libero. Sed sem justo, laoreet vitae, fringilla at, adipiscing ut, nibh. Maecenas non sem quis tortor eleifend fermentum. Etiam id tortor ac mauris porta vulputate. Integer porta neque vitae massa. Maecenas tempus libero a libero posuere dictum. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aenean quis mauris sed elit commodo placerat. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Vivamus rhoncus tincidunt libero. Etiam elementum pretium justo. Vivamus est. Morbi a tellus eget pede tristique commodo. Nulla nisl. Vestibulum sed nisl eu sapien cursus rutrum.

REFERENCES

[1] A. J. Figueredo and P. S. A. Wolf. Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330, 2009.