26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

A LASSO based approach to Wage Spillover

F. Chan [a], M.Harris[a], R.Singh[a] and Wei Ern Yeo[b]

[a]*School of Economics and Finance, Curtin University, GPO BOX U1987, Perth, Western Australia, 6845*
Email: weiern.yeo@postgrad.curtin.edu.au

Abstract: There has been a large amount of literature in the field of Economics on the topic of Wage Spillovers, with many alternate terms such as Key Bargains, Imitation and Wage Leadership. The sheer number of potential crossover effects, and the computational burden presented pose a significant barrier to isolating and determining individual sector effects. This study proposes a novel method wherein regularisation techniques, specifically the LASSO, is applied to this problem in order to both ease computational burden and improve identification and selection of spillover effects. The utility of this method has provided promising results on trial data and this paper aims to eventually test this method on US Wage data encompassing 48 states and 13 industries over a period of 30 years.

Keywords: Shrinkage Estimator, heterogeneous panels, SAR regression

1 INTRODUCTION

The main aim of this paper is to investigate the role that machine learning offers econometric research, mainly through the application of regularisation based techniques to the determination of wage spillovers. To do so, this study will have the three main aims and objectives:

First, an examination of the feasibility of regularisation as variable selection and if and how the process affects the subsequent post-model-selection estimators.

Second, the application of shrinkage estimators to the Econometric field, specifically the Least Absolute Shrinkage and Selection Operator (henceforth referred to as LASSO) examining further the value machine learning can offer Economic studies.

Finally an eventual aim in applying these theories to US labour wage data, which will provide a clearer insight into the spillovers in one of the largest economies in the world.

To achieve these aims, the paper proposes the application of the LASSO to the problem of wage spillovers. But what gap in the wage spillover does the LASSO offer to fill?

2 LITERATURE REVIEW

While there has been a fair amount of papers and research done on the topic of wage spillover, the main gap in the literature this paper hopes to fill is the estimation of more specific spillover effects. For example, Latreille & Manning (2009), in their analysis of spillovers between both unskilled and skilled work and through 88 industries, employs the use of matrices that average wages by their SIC codes. Due to this, while their method does find spillovers between adjacent industries and subgroups, the spillover between industries only measures spillovers from the average of all wages in a matrix, rather than estimating specific spillovers.

Similarly, Driffield & Taylor (2006), uses GMM to estimate spillovers between neighboring regions and related industries, for 19 two digit industries and 11 regions, with an aim to discovering the effect foreign investment has on adjacent wages. In measuring spillovers in the UK, Driffield & Taylor (2006)'s analysis similarly makes use of a matrix to summarise the wages for contiguous regions and industries. Again, while the authors find evidence of wage spillovers between regions and industries, their matrix only gives a broad definition of the spillover, using these matrices to isolate the effect of contiguous regions and industries, rather than isolating the effect of individual contiguous industries or regions.

With this paper's goal of aiming to give some insight on these gaps in mind, the main obstacle that this paper anticipates is the sheer amount of potential "crossover" effects that could surface. For example, while 2 areas yield one possible spillover, 10 areas yield 45. And adding an additional dimension into the mix, a 10 area and 10 industry scenario yields 1485 potential spillovers. Studying these effects would require an immense computational burden and special regression specification, however, perhaps a potential solution for working

out these effects could be provided by regularisation, which is uniquely suited for picking out factors in sparse problems, especially the LASSO proposed by Tibshirani (1996), which this paper will focus on.

3 METHODOLOGY

In order to test the theories presented, this paper proposes an experiment to determine the effect multiple sections such as region, time and industry has on other sections, and, potentially, the effect, if any the crossovers between these sections, such as industry in a region, have on wage change. First, the paper will make using one of the expectation hypothesis wage equations proposed in Turnovsky (1972), specifically, extrapolative hypothesis. The base equation is shown below.

$$\mathbf{w}_t = \frac{\beta_1}{\mathbf{u}_t} + \beta_2 \mathbf{p}_t + \beta_3(\mathbf{p}_t - \mathbf{p}_{t-1})$$

Using this expectations wage equations as a base for wage determination, this paper will then make use of the maximum likelihood Spatial Auto-regressive Regression (SAR) model discussed in LeSage & Pace (n.d.) to take into account the effects that regions have on one another:

$$ln\ \mathbf{L} = -(n/2)\ ln(\pi\sigma^2)\ + ln\ |I_n - \rho W|\ - \frac{e'e}{2\sigma^2}$$
$$\text{e = y - } \rho W y - Z\delta$$

Where y represents a n x 1 matrix of the average annual percent wage change at time t, industry i and time r. Z = $[i_n X]$ where X is a n x K matrix of the explanatory variables from the prior mentioned extrapolative hypothesis, and i_n is a n x 1 vector of 1s associated with the intercept. δ represents a 1 x k+2 vector of coefficients, and W represents weight matrices that display the relation between differing sectors.

In order to perform this regression with the sheer amount of possible weight matrices representing different sectors, this paper aims to subject spillover effects to the LASSO in order to select the ones most relevant to determining the wage, then perform the above Maximum Likelihood regression only using the non-zero LASSO spillovers, plus the . For the LASSO program we will make use of the CVGlasso program for R by Galloway (2018), which makes use of an adaptation of the algorithm proposed by Friedman et al. (2008)

This method is aimed to be tested on 30 years of US wage change data over the 48 continental states (Hence, not including Alaska and Hawaii) and 13 2-digit industrial codes.

4 CONCLUSIONS

This paper has proposed an application of the LASSO regularsation estimator to select spillovers between sectors which hopes to address a gap in the econometrics literature in selecting spillovers and expand on the use of machine learning in Econometrics.

REFERENCES

Driffield, N. & Taylor, K. (2006), 'Wage Spillovers, Inter-regional Effects and the Impact of Inward Investment', *Spatial Economic Analysis* **1**(2), 187–205.
URL: *http://www.tandfonline.com/doi/abs/10.1080/17421770601009825*

Friedman, J., Hastie, T. & Tibshirani, R. (2008), 'Sparse inverse covariance estimation with the graphical lasso', **9**(3), 432–441.
URL: *https://academic.oup.com/biostatistics/article/9/3/432/224260*

Galloway, M. (2018), 'Lasso penalised precision matrix estimation . cvglasso', `https://mattxgalloway.com/CVglasso/`. Accessed: 2025-09-04.

Latreille, P. L. & Manning, N. (2009), 'Inter-industry and inter-occupational wage spillovers in UK manufacturing', *Oxford Bulletin of Economics and Statistics* **62**(1), 83–99.

LeSage, J. & Pace, R. K. (n.d.), *Introduction to Spatial Econometrics*, 0 edn, Chapman and Hall/CRC.
URL: *https://www.taylorfrancis.com/books/9781420064254*

Tibshirani, R. (1996), 'Regression Shrinkage and Selection Via the Lasso', *Journal of the Royal Statistical Society: Series B (Methodological)* **58**(1), 267–288.
URL: *https://onlinelibrary.wiley.com/doi/10.1111/j.2517-6161.1996.tb02080.x*

Turnovsky, S. J. (1972), 'The expectations hypothesis and the aggregate wage equation: Some empirical evidence for canada', **39**(153), 1.
URL: *https://www.jstor.org/stable/10.2307/2552052?origin=crossref*

