DEPARTMENT OF ECONOMICS

A SEMI-PARAMETRIC APPROACH FOR THE TESTING OF DOWNWARD WAGE RIGIDITY USING MICRO-DATA

BY

PARIS NEARCHOU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Department of Economics of the
University of Cyprus

January 2009

ΣΕΛΙΔΑ ΕΓΚΥΡΟΤΗΤΑΣ

Υποψήφιος Διδάκτορας: Πάρης Νεάρχου

Τίτλος Διατριβής: A Semi-Parametric Approach for the Testing of Downward Wage Rigidity using Micro-Data

Η παρούσα Διδακτορική Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για απόκτηση Διδακτορικού Διπλώματος στο Τμήμα Οικονομικών και εγκρίθηκε στις 28 Νοεμβρίου 2008 από τα μέλη της Εξεταστικής Επιτροπής.

Εξεταστική Επιτροπή

Ερευνητικός Σύμβουλος :

- Λούης Χριστόφίδης, Καθηγητής, Πανεπιστήμιο Κύπρου

Άλλα Μέλη :

- Έλενα Ανδρέου, Αναπλ. Καθηγήτρια, Πανεπιστήμιο Κύπρου (Πρόεδρος)
- Άντρος Κούρτελλος, Επίκουρος Καθηγητής, Πανεπιστήμιο Κύπρου
- Patrick Sevestre, Professor, University Paris I Pantheon-Sorbone, France
- Thanasis Stengos, Professor, University of Guelph, Canada

Περίληψη

Η παρούσα διατριβή πραγματεύεται το θέμα της μέτρησης του βαθμού ακαμψίας στους μισθούς. Συγκεκριμένα, έχει ως πρώτο στόχο την ανάπτυξη μιας ημιπαραμετρικής οικονομετρικής μεθόδου που να επιτρέπει την εκτίμηση του βαθμού ακαμψίας προς τα κάτω στους μισθούς, τόσο σε ονομαστικούς όσο και πραγματικούς όρους, η οποία να βασίζεται στη χρήση στοιχείων πάνω στα ποσοστά μεγέθυνσης των ονομαστικών μισθών στο μικροοικονομικό επίπεδο. Ως δεύτερο στόχο έχει την εφαρμογή της μεθόδου αυτής για την εξέταση της ύπαρξης και των δύο τύπων ακαμψίας προς τα κάτω στους μισθούς που συμφωνούνται μέσω συλλογικών διαπραγματεύσεων στον Καναδά, χρησιμοποιώντας στοιχεία από συλλογικές συμβάσεις στη χώρα αυτή. Το περιεχόμενο της διατριβής είναι οργανωμένο σε πέντε κεφάλαια:

Στο Κεφάλαιο 1 παρουσιάζεται μια εισαγωγή όπου σκιαγραφείται το πλαίσιο μέσα στο οποίο εντάσσεται η έρευνα που περιγράφεται στα επόμενα κεφάλαια, καταλήγοντας με την επισκόπηση του υλικού που ακολουθεί.

Στο Κεφάλαιο 2 παρουσιάζεται η επέκταση της μεθοδολογίας 'location-histogram' της Kahn(1997) - η οποία αρχικά σχεδιάστηκε για τον έλεγχο της ύπαρξης του φαινομένου της ακαμψίας προς τα κάτω των ονομαστικών μισθών (DNWR) μόνο - η οποία επιτρέπει ταυτόχρονα και τον έλεγχο για την ύπαρξη ακαμψίας προς τα κάτω στους πραγματικούς μισθούς (DRWR). Επίσης παρουσιάζεται μια πρώτη προσπάθεια φορμαλισμού της προσέγγισης αυτής, και προτείνεται ένας νέος εκτιμητής τύπου FGLS για την εκτίμηση των εξισώσεων παλινδρομήσεων που προκύπτουν κατά την εφαρμογή της. Ακολούθως η εκτεταμένη μεθοδολογία εφαρμόζεται για τον έλεγχο της ύπαρξης και των δύο τύπων ακαμψίας στους μισθούς από συλλογικές συμβάσεις από τον Καναδά, για την περίοδο 1976-1999. Τα αποτελέσματα που προκύπτουν καταδεικνύουν την ταυτόχρονη ύπαρξη και των δύο τύπων ακαμψίας.

Στο Κεφάλαιο 3 προτείνονται προσαρμογές στη μεθοδολογία 'location-histogram' σε σχέση με τον τρόπο ορισμού και εκτίμησης των ιστογραμμάτων πιθανότητας τα οποία χρησιμοποιούνται για να περιγραφούν οι ετήσιες κατανομές των πραγματοποιούμενων μισθολογικών αλλαγών, οι οποίες βελτιώνουν τις θεωρητικές ιδιότητες της μεθοδολογίας αυτής. Ακολούθως επανεκτιμούνται ο βαθμός ύπαρξης των δύο τύπων ακαμψίας για ολόκληρη την περίοδο παρατήρησης, καθώς επίσης και για υποπεριόδους που χαρακτηρίζονται από ομογένεια στο επίπεδο του πληθωρισμού. Τα αποτελέσματα που προκύπτουν για τις υποπεριόδους καταδεικνύουν διαφορετικές μορφές για τους δύο τύπους ακαμψίας ανάλογα με το επίπεδο πληθωρισμού που ισχύει: κατά την περίοδο υψηλού πληθωρισμού μόνο ο τύπος DRWR φαίνεται να ισχύει, ενώ κατά τις περιόδους μέτριου και χαμηλού πληθωρισμού φαίνεται να ισχύουν και οι δύο τύποι ακαμψίας, με τον τύπο DNWR να υπερισχύει εμφανώς στην περίοδο χαμηλού

πληθωρισμού.

Στο Κεφάλαιο 4 γίνεται μια προσπάθεια να αναπτυχθεί ένα ολοκληρωμένο και μαθηματικά αυστηρό υπόβαθρο για την μεθοδολογία 'location-histogram', υιοθετώντας σαν υπόδειγμα περιγραφής των μηχανισμών ακαμψίας που επηρεάζουν τις μισθολογικές αλλαγές μέσα στον πληθυσμό την ημιπαραμετρική εκδοχή του υποδείγματος αναλογικού τύπου ακαμψίας των Goette, Sunde και Bauer (2007): Στην εργασία που παρουσιάζεται, μεταξύ άλλων, γίνεται σύνδεση της μεθοδολογίας 'location-histogram' με τη βιβλιογραφία που αφορά στην εκτίμηση υποδειγμάτων με διακριτές ενδογενείς μεταβλητές, εξηγείται με ποιο τρόπο οι άγνωστες παράμετροι μπορούν επίσης να εκτιμηθούν βάσει της μεθόδου Μεγίστης Πιθανοφάνειας, καθώς επίσης εξετάζονται οι ιδιότητες των εκτιμητών που προκύπτουν από τη μέθοδο αυτή, τόσο θεωρητικά όσο και με τη χρήση προσομοιώσεων. Τα αποτελέσματα από τις προσομοιώσεις δείχνουν ότι η απόδοση του ημιπαραμετρικού εκτιμητή των παραμέτρων των μηχανισμών αναλογικού τύπου ακαμψίας δεν διαφέρει ιδιαίτερα από αυτή του παραμετρικού εκτιμητή που υιοθετούν οι Goette, Sunde και Bauer (2007).

Τέλος, στο Κεφάλαιο 5 παρουσιάζεται μια περίληψη του υλικού που παρουσιάζεται στα προηγούμενα κεφάλαια, και μια σύγκριση των αποτελεσμάτων που περιγράφονται στα Κεφάλαια 2 και 3 με αντίστοιχα αποτελέσματα στη σχετική βιβλιογραφία. Επίσης συζητούνται κάποιες ιδέες για επέκταση της έρευνας που παρουσιάζεται εδώ.

Abstract

This thesis is concerned with the issue of measurement of the extent of downward wage rigidity in wages. Its first aim is to develop a semiparametric econometric method for estimating the degree of downward rigidity in wages in, both, nominal and real terms, that is based on individual level data on nominal wage growth rates. The second aim is to implement the developed methodology to test for the existence of both types of rigidity in the wages agreed through collective bargaining in the Canadian unionised sector, using data from collective agreements in that country. The material presented in the thesis is organised in five chapters:

Chapter 1 provides an introduction; it draws the wider framework that the work presented in the chapters that follow fits into, and concludes with an overview of the material that follows.

In Chapter 2 we present an extension of the 'location-histogram' approach of Kahn (1997) - that was originally designed for the testing of Downward Nominal Wage Rigidity (DNWR) only - that enables the testing for both DNWR and Downward Real Wage Rigidity (DRWR). We also make a first attempt to formalise the approach, and propose a new FGLS estimator for the regression equations that are estimated. The extended methodology is then implement to test for the existence of both types of rigidity in wages from Canadian collective agreements, for the period 1976-1999. The results obtained suggest the presence of DRWR over and above DNWR.

In Chapter 3 we propose modifications to the 'location-histogram' approach with regard to the definition and estimation of the probability histograms that are used to represent the annual actual wage growth distributions, which improve its theoretical properties. We then re-estimate the size of both DNWR and DRWR in the full observation period, as well as obtain estimates within subperiods that are characterised by homogenous inflation levels. The results obtained for the sub-periods suggest different patterns of DNWR and DRWR under different inflation regimes; during high inflation only DRWR is relevant, while both DNWR and DRWR are relevant during medium and low inflation periods, with the importance of DNWR increasing as the inflation level becomes lower.

In Chapter 4 we attempt to provide a comprehensive formal framework for the 'location-histogram' approach, adopting a semiparametric version of the proportional downward wage rigidity model of Goette, Sunde and Bauer (2007) as the underlying wage-setting model: among others, we link this approach with the literature on the estimation of models with discrete endogenous variables, show how the unknown parameters involved could

also be estimated using Maximum Likelihood, and study the properties of the resulting semiparametric estimator both theoretically and using simulations. The simulation results suggest that the performance of the semiparametric estimator of the rigidity parameters of the proportional downward wage rigidity model is not very different from that of the parametric estimator adopted in Goette et al. (2007).

Finally, Chapter 5 concludes with a brief summary of the work presented in the previous chapters, and a comparison of the results obtained in Chapters 2 and 3 with existing results in the literature. Also with discussion of some ideas for extending the work presented here.

To my parents.

Acknowledgments

The work presented here has benefited from my interaction with several people to whom I am grateful:

First, and foremost, I would like to thank Louis Christofides, who has offered me more than is normally expected from his role as my thesis advisor; I will always be grateful to him for his generosity.

Further, I would like to thank Elena Andreou, Yiannis Billias, Andros Kourtellos, and Thanasis Stengos for stimulating discussions that helped me clarify a number of issues pertaining to my work, as well as for suggesting new and interesting paths for my research, some of which are presented here. I must make special reference to discussions with Yiannis Kasparis, which generated some of the core ideas in Chapter 5. Also, I must thank the members of the examination committee for the time and effort they spent to examine my work, and for making helpful suggestions for future extensions.

The work presented here has also benefited a lot from comments received during several presentations of this work at the series of Workshops of Economic Research at the Department of Economics, University of Cyprus. More generally, I am grateful to the Department of Economics for financial support, and for providing the encouragement and support to its research students, including myself, to find their own voice. My participation to the IZA Summer School 2006, in Munich, provided much needed encouragement and inspiration mid-way of my studies. Also the Advanced Summer School in Economics and Econometrics 2008 at Rethymnon, Crete, instructed by William Green, provided a stimulating environment that contributed to the development of the material on Identification that appears in Chapter 5, as well as helped to clarify the bigger picture.

Here I must also thank a few people whose support is reflected in this work despite not having any direct involvement in it: Sofronis Clerides for his positive spirit and generosity since I first arrived in the Department of Economics, Costas Hadjiyiannis for always making me feel to be a member of the team, and Andreas Andreou for consistently providing timely and efficient IT support, which proved critical during the last stages of preparing this work.

I feel that at this moment is also fit for me to express my gratitude to a number of university teachers who, on the way, have inspired me and helped me shape my ethos as an economist, for guidance and encouragement when important decisions had to be taken, as well as for having provided me with support at some of the more difficult moments: Len Gill from the University of Manchester, Andrew Chesher and Simon Burgess from the University of Bristol, Louis Christofides from the University of Guelph, and Panos

Pashardes from the University of Cyprus; I will always be indebted to them.

It is customary in occasions as this one, one to thank one's family for their support; it will be no exception here. My warmest thanks and gratitude to my wife for her patience and understanding during the long process of preparing this thesis, and for the courage I have drawn from her belief in me. And, finally, my thanks and gratitude to my parents for their unwavering support during my studies, and for providing me with all those opportunities, sometimes requiring their own personal sacrifices, to become who I am; this thesis is dedicated to them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The work presented in this thesis is concerned with the search for microeconomic evidence for the existence of downward rigidity in wages. Central to this is the development of a methodology for the joint testing of the hypotheses of Downward Nominal Wage Rigidity (DNWR) and Downward Real Wage Rigidity (DRWR), based on micro-level data on nominal wage growth rates. Further, the implementation of this methodology to test for the existence of both types of rigidity in the wages agreed through collective bargaining in the Canadian unionised sector.

The study of wage rigidity, and in particular, downward wage rigidity, has a long tradition in economics. It goes back, at least, as far as to Keynes (1936), who placed DNWR at the centre of the theory he proposed to explain the existence of involuntary unemployment. Since then, a significant amount of research effort has been directed towards understanding issues relating to wage rigidity, both theoretically and empirically. Goette et al. (2007) summarise the issues that draw considerable attention still today, in the following three questions: Firstly, 'what are the *consequences* of downward wage rigidity', secondly, 'what are its *causes*', and, thirdly, 'what is the extent of rigidity in practice' (*measure*).

Since the mid 1990's research in this field has gained new impetus, taking at the same time new directions. Most notably towards the micro-level: In the theory area, examining the micro-foundations of the theories that link DNWR to involuntary unemployment, and the microeconomic mechanisms that produce DNWR. Empirically, looking for direct microeconomic evidence for the existence of both DNWR and DRWR using data on individual nominal wage growth rates.

The aim of this introductory chapter is to draw the wider framework that the work presented in this thesis fits into, and also to give a picture of the material that follows. In Section 1.1 we provide a brief overview of the developments in the literature in the three areas of interest mentioned above, including some discussion on policy issues. Then, in Section 1.2, we provide an overview of the three main chapters that follow.

1.1 Downward Wage Rigidity

1.1.1 Definitions, Causes, and Consequences

In a market economy, prices help determine the allocation of resources among competing uses. When the demand for, or supply of, a good shifts because of changes in tastes or technology, relative prices adjust, allowing resources to be reallocated to reflect these changes.

This process of adjustment to fundamental changes in the economy also applies to labour markets through changes in real wages. For example, changing consumer tastes or technological innovation may cause some forms of labour to become relatively more valuable than others. Then adjustments in the relative real wages can induce labour market flows that lead to a new equilibrium in the labour market, one that accommodates the new fundamentals. From the above it follows that if, for some reason, the real wages are rigid,[1] then the market mechanism may not work efficiently.

Despite its importance, *labour market inefficiency* is not the main reason for the interest attracted by DNWR,[2] a type of wage rigidity. Rather, it is the macroeconomic implications of DNWR in situations where underlying changes in supply and demand conditions create pressures for decreases in the real value of wages at individual firms or in certain sectors of the economy; under such conditions it is argued that DNWR could induce *involuntary unemployment*. To see how, we note that with moderate rates of inflation, any required decreases in the real wages can be achieved by letting actual (nominal) wages increase by less than the rate of inflation. With zero inflation, however, real wages can be reduced only by cutting nominal wages. The presence of DNWR would prevent such nominal cuts from being realised, and real wages would remain at a level higher than the market clearing level.[3] Then employment flows could substitute for wage flexibility, resulting in unemployment.

There has been considerable discussion recently about the ability of inflation to facilitate the adjustment of real wages and so enhance the performance of the economy in such situations: If nominal wages are downwardly rigid, then the only way that real wages can be reduced is to allow inflation to erode the real value of nominal wages over time - the higher the rate of inflation, the faster the buying power of downwardly rigid nominal wages is eroded. The idea that moderate levels of inflation can 'grease the wheels of the labour market' can be traced back to Keynes, and was further developed by Tobin (1972). Recently Akerlof, Dickens and Perry (1996) have gone a step further, suggesting that there exists a long-run trade-off between inflation and unemployment - that is, a long-run Philips-curve-type relationship. This result implies that it might be optimal for Central Banks to aim for moderate inflation levels in the long-run rather than pursue a zero-inflation policy, as the benefits from moving from a near-zero to moderate inflation levels, in terms of reduced unemployment, could outweigh the costs associated with higher inflation.

Further to the macroeconomic consequences of wage rigidity discussed above, another issue relates to its implications for *inflation dynamics*. In particular, it has been suggested that rigid wages can be a cause of less frequent changes in prices of products with a high labour share. In turn, price stickiness could leads to higher output volatility in response

[1]Wage rigidity describes the situation where the wage does not adjust but remains fixed at its current level. A related phenomenon is wage stickiness, which refers to the situation where adjustments can be achieved only very gradually.

[2]See Groshen and Schweitzer (1996).

[3]Therefore in a low inflation environment, DNWR induces DRWR.

to shocks, which requires stronger interest rate changes to affect inflation.[4]

While DNWR has important microeconomic and macroeconomic implications, one of the long-standing puzzles of labour economics is why don't wages fall during recessions. The theoretical literature has identified a number of possible mechanisms that render wages rigid, in either real or nominal terms:

With respect to mechanisms that produce DRWR, one branch of the literature focuses on the role of unions and collective bargaining agreements in generating consistently higher wages than the market clearing level; see for instance, Layard, Nickell and Jackman (1991). Such *insider-outsider* theories also imply that wages are not responsive to changes in the economic environment.

Another strand of literature examines under what conditions firms have an incentive to pay wages above the market clearing level (*efficiency wage theories*). For example, in Shapiro and Stiglitz (1984), or Macleod and Malcomson (1993), firms pay high wages so that workers would lose a quasi-rent if fired. Workers therefore have an incentive to provide high effort, even when effort is observable but not contractible. In a variant of these models (Akerlof and Yellen (1990)), firms pay high wages because workers are reciprocal and respond to higher wages by exerting higher effort (*reciprocity model*). These theories imply that real wages are rigid in the sense that they cannot easily be adjusted, in particular downwards, because of the incentives of firms to pay high wages. Therefore all the models above predict involuntary unemployment.

With respect to mechanisms that produce DNWR, it has been a long-standing conjecture that individuals, as well as being concerned about changes in their real wage, they also care about changes in their nominal wage (Tobin (1972)). More recently, work by economists and psychologists has made this intuition more precise, providing empirical evidence based on employer surveys and laboratory experiments. This evidence suggests that fairness judgments - made by employees - about wage contracts are made relative to a reference transaction (Kahneman, Knetsch and Thaler (1986)). Furthermore, this reference point seems to be the current nominal contract in many cases (Shafir, Diamond and Tversky (1997)). Nominal wage cuts are therefore perceived as particularly unfair, thus affecting morale, which can lead to sharp reductions in effort.[5]

A common criticism of this theory is that it implies money illusion, as employees fail to appreciate the implications of inflation on the purchasing power of their wages, and is, thus, inconsistent with rational behaviour. Holden (1994,2004), on the other hand, has shown that when unions and workers have to bargain under the so-called holdout rule, i.e., that a contract can only be changed by mutual consent, nominal wage rigidity can be the equilibrium outcome even if all actors only care about real wages.

Finally, another source of DNWR that has been cited is the existence of institutional

[4]See, for example, Fougère, Le Bihan and Sevestre (2005), Altissimo et. al (2006), Álvarez et. al (2006), Dhyne et. al. (2006), and Vermeulen et. al. (2007).

[5]Bewley (2004) provides a survey of studies in both the economics and psychology fields. Overall, these provide consistent support for the fairness argument.

characteristics, such as legal restrictions regarding a legislated minimum wage.

1.1.2 Measure

Until the mid ninety's, the predominant approach towards testing for the existence of DNWR was to test for counter-cyclicality in the aggregate real wage level. This is an indirect approach; it is based on the prediction of macro models that real wages should be counter-cyclical, that is, they should be relatively high in recessions and low in expansions.[6]

A fundamental shift in the overall approach towards measuring the extent of downward wage rigidity took place with the work by McLaughlin (1994). He was the first to look for direct evidence for the existence of DNWR in micro-level data on nominal wage growth rates, using the properties of the actual wage growth rates distribution (WGD) when DNWR is present.[7] Since then a considerable number of studies have been carried-out following this microeconometric approach, covering most of the developed countries,[8] based primarily on data on individual employee wage growth rates[9] (from panel surveys as well as administrative panel data). Initially efforts were concentrated on testing for the existence of DNWR only.[10] However, since the early 2000's, the attention has shifted towards testing for both types of rigidity.[11]

This new testing methodology is based on the idea that the presence of each type of downward wage rigidity introduces its own distinct type of distortions to the shape of the actual WGD, therefore, a test for the presence of each type of rigidity could be based on detecting these particular types of distortions. Broadly speaking, there have been two main empirical approaches of designing tests along these lines; one of comparing the shapes of the distributions of the actual (factual distribution) and rigidity-free wage growth rates (counterfactual distribution),[12] and the other of testing for attributes of the shape of the actual WGD, such as for (the lack of) symmetry.

The practical implementation of this type of tests poses two main challenges;

Firstly, the shape of the *actual* WGD is estimated using data on *observed* wage growth

[6] See Groshen and Schweitzer (1996) for a brief review of alternative macroeconometric test approaches.

[7] Practical reasons could also account for the timing of this shift: the availability of extensive and reliable panel data sets, as well as the lowering of the cost of computer power that made the highly computer intensive calculations feasible.

[8] More than often these studies were either sponsored or directly performed by the central banks in these countries.

[9] Also, there are examples of the use of data from collective agreements (e.g.Crawford (2001), Christofides and Stengos (2003), and Christofides and Leung (2003)), and industry level data (Holden and Wulfsberg (2007)).

[10] And, also, menu costs.

[11] This was justified by the fact that the distinction between real and nominal wage rigidities is important beyond the purely conceptual level, as the source of rigidity is also important for economic policy. In the presence of real rigidity, monetary policy should primarily aim at stable prices, implying procyclical adjustments of money supply (see Goodfriend and King (1997)). On the other hand, as already discussed, if nominal wages are downwardly rigid, a positive rate of inflation may 'grease the wheels of the labour market', which would rather ask for monetary policy to counteract demand shocks (Tobin (1972), Akerlof et al. (1996)).

[12] Typically, by comparing the amount of probability mass that is allocated in that part of the actual distribution that is expected to be distorted.

rates. However, such data often suffer from measurement error, especially if they originate from non-administrative panel surveys. Consequently, the estimates of the actual WGD obtained in this way are not precise. When this problem is not ignored, it is either tackled by accounting explicitly for the presence of measurement error in the model for the data generating process of the observed wage growth rates (see Altonji and Devereux (2000)), or by introducing a preliminary stage of 'correcting' the observed wage growth rates for measurement error, before proceeding with the implementation of the test (see Dickens and Goette (2006)).

Secondly, the distribution of rigidity-free, or 'notional', wage growth rates is, typically, unknown. Furthermore, it cannot be directly estimated as the notional wage growth rates are unobservable. This issue is addressed by using data on *observed/actual* wage growth rates to infer the shape of the *notional* WGD. This requires making some additional assumptions about the shape of the notional WGD. In the literature, a variety of such sets of identifying assumptions have been employed, which are characterised by different degrees of parameterisation of the notional WGD.[13] In the case of the testing approaches that deal only with the presence of DNWR, examples include: Altonji and Devereux (2000) ('earnings-function' approach - parametric, accounts for measurement error), Kahn (1997) ('histogram-location' approach - semiparametric), Card and Hyslop (1997) ('symmetry' approach - nonparametric), McLaughlin (1994) and McLaughlin (1999) ('skewness-location' approach - parametric/semiparametric, depending on the measure of skewness), Christofides and Stengos (2001) (symmetry test - nonparametric), Knoppik (2007) ('kernel-location' approach - nonparametric).[14]

Not all of the approaches proposed for the testing of DNWR can be extended to allow the testing of both DNWR and DRWR. In fact, only the 'earnings-function' approach has been extended accordingly, and this is discussed in Goette et al. (2007). An alternative approach is the one developed by the International Wage Flexibility Project (IWFP), which is also parametric; see Dickens and Goette (2006).[15]

1.2 Thesis Overview

The work presented here fits into the latter strand of the literature discussed above, the one concerned with the measurement of downward wage rigidity. In a nutshell, it describes

[13] It has also been noted in the literature, that these testing approaches are, in effect, statistical in nature; at best their set-up only makes an indirect reference to the theoretical models that explain the emergence of DNWR and DRWR. See Holden (2002), and also Kramarz (2001) for discussion on this.

[14] With respect to the results obtained from these and other studies that follow the same testing approaches, the empirical work on DNWR produced mixed results, even for the same countries (see review in Kramarz (2001) and Palenzuela, Camba-Mendez and Garcia (2003)). This was attributed to the diversity of methodologies used to produce these results. A significant attempt to reconcile the results was undertaken by the International Wage Flexibility Project (IWFP), which also extended the scope of the analysis to include also DRWR. The results from this effort are reported in Dickens et. al. (2007).

[15] Some of the above studies have also attempted to identify the causes of rigidity; see, among others, Dickens et. al. (2007) and Du Caju (2008).

a formal framework within which the original 'location-histogram' approach[16] proposed by Kahn (1997), that was intended for the testing of the existence of DNWR only, is extended to allow the testing for both DNWR and DRWR, and provides a number of analytical and simulation results regarding its properties. Furthermore, the extended methodology is implemented to test for the presence of both types of rigidity in wages from collective agreements in the unionised sector of Canada.

The data used have been provided by the Human Resources Department Canada, the federal ministry in Canada responsible for monitoring agreements between firms and unions. This data set presents certain characteristics that render it particularly suitable for this type of investigation. Most importantly, it covers a relatively long period, from 1976 to 1999, during which Canada experienced a wide range of inflation levels, from as high as 12.43% in 1982, to as low as 0.16% in 1995. As the level of inflation in a particular period determines, to a great extent, which part of the actual wage growth distribution is distorted by either type of downward wage rigidity, having this kind of variation makes it possible to identify the effect of both types of rigidity, if present.

In addition, this data has been used in the past by other researchers to investigate the extent of downward wage rigidity. In particular, Christofides and Leung (2003), using a variant of the original 'location-histogram' approach, have found clear evidence of DNWR. Also, the same conclusion in favour of the presence of DNWR was reached by Christofides and Stengos (2002) and Christofides and Stengos (2003), who used a different methodology to detect DNWR. Therefore, our work can provide a robustness check for this result when one allows for the presence of DRWR as well. At the same time, our empirical work will provide the opportunity to verify the conclusion reached by Christofides and Li (2005), who used a different methodology than the one used here, in favour of the presence of both DNWR and DRWR. With regard to the presence of DRWR, we would expect that this is more likely to be present in wages agreed through collective bargaining rather than in non-collectively bargained wages, as unions are expected to be particularly concerned with preserving, at least, the real value of wages. The fact that the data analysed here concerns agreements with relatively wide coverage - involving unions that cover more than

[16] The 'location-histogram' approach of Kahn uses changes in the shape of the wage change distributions to identify the counterfactual distribution. It consists of two steps:

First, annual relative-frequency histograms for the observed wage growth rates are constructed in order to compute factual bin heights. The location of the bins in the annual histograms is determined relative to the location of the sample median in each year.

Then the method proceeds to compare the height of the bins - located at a given distance from the median - in years when rigidity may be binding (when those bars fall below zero) with their height in years where rigidities are absent (when they fall above zero). This is done by regressing the factual bin heights on dummies indicating the position of the bins relative to zero; in its simplest specification, a dummy indicating bins with negative wage growth values, and a dummy for (around) zero wage growth values. The larger the positive coefficient on the dummy for zero wage changes, the higher the spike at zero, i.e. the larger the proportion of wage freezes. The larger the negative coefficient on the dummy for negative wage changes, the higher the proportion of wage changes that have not been cut. Taken together, these provide evidence of downward wage rigidity.

The method is semiparametric in nature, as it requires no assumption on the distribution of wage changes, although it does makes assumptions about the functional form of the distortions attributed to rigidity. except from its simplest specification.

200 employees, - which suggests relatively high bargaining power for the unions, makes this result more likely to emerge from our empirical analysis.

A final point regarding the positive features of the data analysed here, is that, as it is administrative in nature, this means that measurement error in the observed wage growth rates is not an issue, and thus there is no need to allow for this in the modeling exercise.

With regarding to possible limitations of the analysis presented here, first we note that any results obtained refer directly to, at most, a relatively small part of the workforce in Canada (roughly 11%), and possibly, their generalisation to the non-collective wage agreements may not be appropriate. Also, the wage growth data analysed here refer to the basic wage rather than to full earnings, that typically include addition financial and non-financial benefits for the workers. Therefore, despite being important as a result on its own right, any rigidity found in this remuneration measure may non reflect the flexibility actually present in total remuneration. Finally, the analysis presented here does not look for possible determinants of, or more generally, the sources of heterogeneity for downward wage rigidity, such as the duration of the wage agreements, and the industry the employer operates in, as information on such characteristics were not available to us at the time of the analysis.

The exposition of the material presented in this thesis is organised in three chapters:

1.2.1 Real and Nominal Wage Rigidities in Collective Agreements

In this Chapter we consider an extension of the original 'location-histogram' approach to enable the testing for both DNWR and DRWR. We also make a first attempt to formalise this approach. In particular we take the view that the annual relative-frequency histograms constructed in stage 1 of the approach can be interpreted as estimates of probability histograms corresponding to the underlying (continuous) wage growth distributions. We also derive the functional form of the variance-covariance of the errors in the regression model (stage 2), which we then use to apply the FGLS estimator, and also calculate the corrections for the standard errors of the OLS estimator.

The extended methodology is implement to test for the existence of both types of rigidity in wages from Canadian collective agreements, for the period 1976-1999. This data set is suitable for this type of analysis, as the observation period covers years with diverse inflation experience, which is necessary for the identification of the distortions produced by both types of rigidity. This part builds upon and extends the work by Christofides and Leung (2003), who use a variant of the 'location-histogram' approach to test for DNWR using the same data set.

1.2.2 Patterns of Nominal and Real Wage Rigidity

The work presented in this Chapter considers extensions and improvements to the work discussed in Chapter 2. The main are the following:

1. We propose an alternative estimator for the probability histograms of the nominal-wage-growth distributions (in stage 1 of the estimation procedure) that is based on a Kernel estimator of the cumulative distribution function, that overcomes some of the weaknesses of the relative frequency estimator used in Chapter 2.

2. We adopt an alternative 'standardisation' rule - from the one used in Chapter 2 - to construct the probability histograms, that is more commonly used in the literature. This enables us to check the sensitivity of the results of Chapter 2 on the adopted standardisation rule.

3. We test for the presence of both types of rigidity separately in three sub-periods of the full period covered by the data, which are characterised by homogenous inflation levels. We are thus able to trim the model fitted to the whole observation period in Chapter 2 to the specific features of these sub-periods, and check whether the results obtained for the whole period are confirmed for the sub-periods. Furthermore to examine the - possibly changing - patterns of DNWR and DRWR as the level of inflation changes.

1.2.3 A Semiparametric Approach for the Testing of Proportional DWR

The work presented in this Chapter aims towards providing a better understanding of the methodology developed and implemented in Chapters 2 and 3.[17]

On the one hand we make clarifications to the formal framework introduced in Chapter 2, making an effort to carefully outline all the assumptions that underly this methodology, and discussing their implications. The insights gained from this exercise then enable us to distinguish the modeling aspect of the approach from the estimation method. In particular, we take the view that the modeling aspect is concerned with the approximation of the continuous notional WGD with a discrete distribution, and the specification of a model for the probability mass function of the latter - i.e. the probability histograms - that allows for the presence of distortions consistent with the existence of the two types of downward wage rigidity. Adopting this perspective is particularly useful as it allows us to borrow results from the literature on the estimation of models with discrete endogenous variables, i.e. 'discrete choice models', to clarify issues on the estimation aspect of the 'location-histogram' approach. Specifically, we are able to show how Maximum Likelihood can be

[17] Because this work historically followed the work presented in Chapters 2 and 3, some of the results presented here are not embodied in that work. Also, the way the implementation of the methodology is organised into stages is slightly different here from what it is in these chapters, however there are no differences in substance.

used to estimate the size of the distortions in the probability histograms. We also show that this is asymptotically equivalent to the regression approach of Kahn (1997), which can be viewed as an application of the Minimum Chi-Squared Estimation method for multivariate discrete endogenous variables discussed in Zellner and Lee (1965).

To perform this analysis we used as an underlying wage-setting model the proportional Downward Wage Rigidity (DWR) model, similar to the one considered by Goette et al. (2007), except from that, for our purposes this was specified in a semi-parametric form. The reasons for the choice of this model were: (a) that it provides a generalised framework to work with, as the likelihood function that is maximised to obtain the estimates of the rigidity parameters is built up the from the very basic level, that is, from the relationship between the notional and actual wage growth rates, and (b) the model for the discrete approximation of the continuous actual WGD that emerges includes the original parameters of the rigidity mechanism; demonstrating how these could be estimated semiparametrically can be interesting on its own right as this is complementary to the work of Goette et al. (2007) who estimate them using a parametric approach.[18]

The Chapter concludes with the discussion of simulation results that aim, on the one hand, to investigate certain properties of the approach, and on the other hand, to compare its performance with that of the parametric ('earnings-function') approach used by Goette et al. (2007).

[18]We note that the proportional DWR model described here implies different type of distortions than those specified and estimated in Chapters 2 and 3. Although this model could have also been fitted to the data used there, this was beyond the scope of the work reported here.

Chapter 2

Real and Nominal Wage Rigidities in Collective Agreements

2.1 Introduction

Monetary policies in a number of countries have, at least until the current oil price shocks, succeeded in limiting price inflation. A by-product of this success has been concern with the extent to which this inflation record has been achieved at a cost. In a low inflation environment, downward nominal wage rigidity (DNWR) may mean that nominal-wage reductions, called for by bargaining pair-specific productivity shocks, do not occur, thereby compromising the efficiency of the labour market. Indeed, some studies go as far as to look for the unemployment consequences of such low-inflation mechanisms. If inflation greases the wheels of the labor market, then its absence may lead to costs. An expanding literature covering a number of countries takes advantage of the recent periods of low price inflation and attempts to measure the extent and consequences of DNWR.[1] This literature has been further energized by the International Wage Flexibility Project (IWFP), led by William Dickens and Erica Groshen.[2]

An important concern of studies in this literature should be the extent to which real rigidities can be treated as part and parcel of the more general wage adjustment process. Naturally, the extent to which price inflation and particularly anticipated price inflation feed into nominal-wage adjustment is a subject that goes at least as far back as Friedman (1968). While nominal-wage adjustment is clearly conditioned by price inflation effects, the extent to which downward real wage rigidity (DRWR) exists, its implied impact on the shape of the wage adjustment distribution in the neighborhood of the anticipated rate of inflation, and possible interactions of this process with DNWR are issues that deserve further attention.

A particularly good data set for studying these effects is the Human Resources Development Canada (HRDC) record of the provisions of collective bargaining agreements reached in the Canadian unionised sector. The data is thought to be very accurate because it refers to legally binding provisions, it covers all industries over all of Canada, and it covers high as well as low inflation periods since 1976. In an earlier paper by Christofides and Leung (2003), the HRDC data were used to examine DNWR and menu cost behaviour in the period 1976-1999 using parametric techniques inspired by Kahn (1997). In this paper,

[1] An extensive review of the literature is contained in Christofides and Leung (2003).

[2] Much more information is provided in the proceedings of the project's Final Conference in Dickens and Groshen (2004).

we extend the earlier study to more explicitly encompass DRWR and its interaction with DNWR. A strength of the HRDC data for current purposes is that the diverse inflation experience that it encompasses makes it possible to differentiate DNWR from DRWR processes. The results obtained indicate significant and substantial nominal and real wage rigidity in the contract data.

Our approach is distinctively different from other recent studies that also test for the presence of both types of rigidity, including, among others, those of Bauer, Bonin, Goette and Sunde (2007) for Germany and Barwell and Schweitzer (2007) for the UK. Both studies find evidence for the presence of both types of rigidity, with real rigidity being more pervasive. Their approach builds upon the maximum-likelihood methodology originally proposed by Altonji and Devereux (2000) for the testing of DNWR alone, and requires parametric assumptions about the family of the rigidity-free nominal-wage-growth distribution. In contrast to them, we make no such assumptions, nor do we impose a symmetric structure, as it has often been done in studies that examine the presence of DNWR. For the identification of the shape of the rigidity-free distribution and the size of the distortions due to the presence of rigidity we exploit the fact that we have several yearly samples from nominal-wage-growth distributions whose shape is affected by Downward Wage Rigidity (DWR) differently from year to year.

The rest of the paper is organised as follows: In Section 2.2, we consider the effect of the presence of each type of rigidity on the wage-growth distribution and in Section 2.3 we present more details on the data and sources. The empirical specification and estimation issues are presented in Sections 2.4 and 2.5 respectively. The results obtained are described in Section 2.6, and concluding observations appear in Section 2.7.

2.2 Downward Wage Rigidity and Wage Growth Distributions

We take DNWR to describe that feature of the wage adjustment process where agents, individual employees or unions, are reluctant to accept a nominal-wage cut (negative wage growth) and instead would settle for a nominal-wage freeze (zero growth). Justifications for nominal rigidity range from the comparability and fairness arguments documented in Bewley (1999) to the theoretical papers by Macleod and Malcomson (1993), Malcomson (1997), Holden (1994) and Holden (2004) which build on the notion that nominal wages can be changed only by mutual consent.

At the population level, this reluctance would mean fewer cuts in nominal wages and more nominal-wage freezes *relative to* the case of no rigidity. In terms of the distribution of nominal-wage-growth rates, this translates into a shift of probability mass from negative values of the support of the distribution towards the point zero. Therefore the rigidity-contaminated nominal-wage-growth distribution would show a deficit of probability mass for negative values of the support, and a surplus at point zero, relative to the rigidity-free (or notional) distribution. At the same time, the two distributions would be identical beyond the point zero.

We can see these effects in the two diagrams of the top row of Figure 2.1, where we have simulated the rigidity-free and the rigidity-contaminated nominal-wage-growth distributions for two particular types of DNWR mechanisms. In these two diagrams, as well as the rest in Figure 2.1, both distributions are represented by their probability histograms. The horizontal axis measures the nominal-wage growth, in percentage rates, and the vertical axis the probability mass that falls in the bins of the histogram. The rigidity-free distribution is always represented by the light-shaded bars, and the rigidity-contaminated by the dark-shaded bars. In these diagrams, the distortion in the nominal-wage-growth distribution is manifested by the difference between the height of the corresponding bars of the rigidity-contaminated and notional histograms. The distortion could take the form of a (probability mass) deficit when the bars of the rigidity-contaminated histogram are shorter than the corresponding bars of notional histograms, and the form of a surplus when the opposite is true. An overall shift of probability mass to the right is detected when there is a collection of bins with a surplus that lies to the right of a collection of bins with a deficit.[3]

In the leftmost diagram, we consider the case of *absolute* DNWR, where all agents facing a nominal-wage cut succeed in settling for a nominal-wage freeze. Therefore, at the population level, there should be no nominal-wage cuts, but instead an 'excessive' amount of nominal-wage freezes. In the diagram we see that, although in the absence of rigidity there would be a number of wage cuts, indicated by the positive height of the bars of the notional probability histogram for the bins with negative values, the height of the corresponding bars of the rigidity-contaminated histogram is zero. At the same time, for the latter histogram, the missing mass is concentrated in the bin that contains the point zero, while the two histograms coincide beyond that. The case of absolute DNWR could be considered as the extreme scenario of *partial* DNWR. At the population level, the case of partial DNWR (rightmost diagram, row 1, Figure 2.1) would mean that there is a positive number of nominal-wage cuts, but also that there is an 'excessive' amount of nominal-wage freezes, although to a lesser extent than the case of absolute DNWR, other things being equal.[4]

DRWR can be defined in a similar way to DNWR. In particular, it is taken to describe the situation where agents are reluctant to accept real-wage cuts but instead would settle for a real-wage freeze. In practice, this attitude takes the form of reluctance towards accepting reductions in the *anticipated* real wage since, at the time of bargaining, future

[3]In order to make the comparison of the effects of the various types of wage rigidity easier, we have drawn the same rigidity-free distribution in all diagrams. We have also deliberately made the bins of the rigidity-contaminated histogram narrower in order to be easier to distinguish the two histograms.

[4]In all simulations presented in this paper where we look into cases of partial downward wage rigidity, we have assumed the 'proportional' type, where each agent facing a wage cut, either in nominal or real terms, depending on the type of rigidity we examine, faces the same probability of settling for the respective type of wage freeze. This assumption does not influence our conclusions.

inflation is typically unknown. The anticipated real-wage level is based on their belief, at the time of bargaining, about the future level of inflation and might be determined as described in theoretical constructs such as efficiency wages, efficient bargains, and implicit contracts. In this paper, we do not concern ourselves with how this anticipated real wage might be determined.

As in the case of DNWR, the presence of DRWR would distort the shape of the nominal-wage-growth distribution. To see how this could happen, we first note that DRWR could also be described as the situation where agents are reluctant to accept nominal-wage-growth rates that are below their anticipated rate of inflation for the period the wage is bargained for, and instead would settle for growth rates that are equal to that. At the population level, this would mean that agents who face nominal-wage growth at a rate below anticipated inflation would settle for a nominal-wage increase equal to the anticipated rate of inflation. Consequently, the presence of DRWR would shift probability mass to the right, from smaller values of nominal-wage growth towards the values of anticipated inflation in the population. The exact form of the shift of mass to the right towards the values of anticipated inflation depends on the nature of the rigidity mechanism and the joint distribution of the notional (nominal) wage growth and anticipated inflation among all agents.

Nevertheless, without any distributional assumptions, it is possible to distinguish three regions in the nominal-wage-growth distribution for which we can make qualitative predictions about the nature of the distortions introduced. Firstly, the interval of values that lies to the left of the support of the distribution of anticipated inflation, if one exists, could only loose mass to the right since all agents whose nominal-wage growth falls in this region face the prospect of a real-wage cut. Therefore, in this region, the rigidity-contaminated distribution can only exhibit a deficit. Secondly, the interval of values that lies to the right of the support of the distribution of anticipated inflation, if one exists, would not be distorted, since all agents whose nominal-wage growth falls in this region face the prospect of a real-wage increase. Finally, the interval of values that corresponds to the support of the distribution of anticipated inflation, will attract mass from its left, and therefore for this interval the rigidity-contaminated distribution will exhibit a surplus, *in total*. However, it is possible that, in some parts of this interval, the rigidity contaminated distribution will exhibit a deficit. In terms of the probability histogram, we can understand how this could happen by noting that a particular bin that contains values of anticipated inflation can attract mass from bins to its left but at the same time loose mass to bins to its right that also contain values of anticipated inflation. The net effect cannot be clear without knowledge of how the notional-wage growth and anticipated inflation are jointly distributed. The only exception is the rightmost bin in this region, for which we know that it cannot exhibit a deficit since all other bins that contain values of anticipated inflation lie to its left. Despite this uncertainty, we could assume that it would be more likely that bins that lie further to the left in this interval will show a deficit and bins further to the right will show a surplus.

In order to see how in practice DRWR could distort the shape of the nominal-wage-growth distribution, we consider several examples of the presence of DRWR that differ from each other with respect to the characteristics of the distribution of anticipated inflation and the extent of the rigidity.[5] First, we consider the case of *firm and uniform beliefs*, where all agents anticipate the same value of future inflation (marked by the vertical broken line in the diagrams of row 2, Figure 2.1). In the case of *absolute* DRWR, the presence of DRWR will shift probability mass towards the value of anticipated inflation from its left, in a similar way that the presence of DNWR will shift probability mass from negative values towards the point of zero nominal-wage increase (leftmost diagram, row 2, Figure 2.1). In the case of *partial* DRWR, there remain a number of anticipated real-wage cuts but there is also an 'excessive' number of wage increases just equal to the anticipated rate of inflation (rightmost diagram, row 2, Figure 2.1).

More realistically, agents disagree on their beliefs of the future level of inflation, and thus there is a distribution of values of anticipated inflation among the population members.[6] First, we consider as a benchmark the implausible scenario where the marginal distributions of notional-wage growth and anticipated inflation coincide on a point-by-point basis; that is, each and every agent faces a notional-wage growth exactly equal to the agent's level of anticipated inflation. In this case, all agents effectively experience an anticipated real-wage freeze, and, therefore the actual distribution will coincide with the notional. If we relax the assumption that all agents face the prospect of an anticipated real-wage freeze and, instead, assume that some are faced with an anticipated real-wage reduction (and at the same time keep the assumption that the two marginal distributions coincide), then the presence of DRWR could shift some probability mass to the right (leftmost diagram, row 3, Figure 2.1).[7]

More likely, the two marginal distributions will not coincide, the support of the anticipated-inflation distribution can be assumed to lie within the support of the nominal-wage change distribution, and there will be agents in the population that face the prospect of an anticipated real-wage decrease. Examples of this case are depicted in the rightmost diagram of the third row, and the leftmost diagram of the fourth row of Figure 2.1. In the first diagram we consider the case where all the agents who face the prospect of a real-wage cut manage to avoid it (diagram with caption Absolute DRWR(b)), while in the second diagram only some agents manage to avoid it (diagram with caption Partial DRWR(b)). The nature of the distortions in the two cases is qualitatively the same but they differ in terms of the size. Also note the nature of the distortion in the three regions of the rigidity-contaminated distribution relative to their position to the support of anticipated inflation (i.e. the absence of mass to the left of the minimum of the support of anticipated-inflation distribution, the absence of any effects to the right of its maximum, and the diverse patterns within it).

It is interesting to see what the presence of DRWR means for the distribution of *actual* real-wage growth. If we accept that typically the distribution of anticipated inflation ex-

[5]In all simulations we assume that nominal-wage growth and inflation beliefs are independent. This does not affect our conclusions.

[6]Firm and uniform beliefs could be seen as a special case of this.

[7]In this diagram, as well as the remaining diagrams in Figure 2.1, the probability density function of the non-degenerate anticipated-inflation distribution is depicted by the solid thick line.

tends below and above the realised inflation value, then the presence of DRWR is consistent with observing real-wage cuts (relative to the realised value of inflation), even in the case of absolute DRWR. Therefore, the occurrence of real-wage cuts does not, in general, suggest that DRWR does not exist; real-wage cuts would, however, rule out the case of absolute DRWR and perfect foresight.

Finally, when some collective agreements are affected by DNWR and others by DRWR, then both types of distortions will be present in the shape of the actual-wage-growth distribution. This case is depicted in the rightmost graph of the bottom row of Figure 2.1, where there is both a spike at the bin containing the point zero and deficit in probability mass for bins to the left as well as to the right of that bin. Note that the two types of distortions have similar effect at the bins below zero, i.e. they reduce the probability mass concentrated there. On the other hand, they have opposite effects at the bin containing zero, since the presence of DRWR shifts mass from that bin to other bins to its right (negative effect), while the presence of DNWR shifts mass to that bin from bins to its left (positive effect). The nature of the combined effect will depend on the proportion of agreements affected by each type of rigidity, as well as the intensity of each type. Moreover, there is probability surplus for the bins that lie towards the right tail of the distribution of anticipated inflation and no effect to the bins that lie beyond the maximum level of anticipated inflation.

2.3 Data

The contract data used in this paper are compiled by HRDC, the federal ministry responsible for monitoring agreements between firms and unions. The database[8] contains information on provisions for 10,945 wage contracts signed in the Canadian unionised sector and involves settlement dates as early as 1976 and as late as 1999. The agreements cover bargaining units involving 200 to nearly 80,000 employees, in both the private and the public sector, and their duration ranges from a few months to several years. Because reporting requirements apply, this information is thought to be very accurate. The data set that is used for the empirical analysis contains one observation for each contract, and the recorded information is the growth rate of the total nominal-wage adjustment $(WNC + COLA)$[9] over the whole of the life of the contract, calculated in *annual* terms. The observation for each contract is allocated to the year the contract became effective.

Table 2.1 shows the number of contracts and the sample mean, median and standard deviation of the nominal-wage growth rate for each year in the observation period.[10] Also,

[8]See Christofides and Stengos (2003) for a detailed description.

[9]Our analysis deals with the total wage adjustment, which is composed of the non-contingent wage adjustment (WNC) and $COLA$. It should be noted, however, that, because the incidence and intensity of COLA clauses is limited throughout the observation period, the results we obtain are similar to those that are based on the analysis of the non-contingent wage adjustment data.

[10]Because of the smaller number of contracts, the first two and the last three years in the sample are considered together in everything that follows.

the corresponding annual rate of Consumer Price Index inflation (CPI) and an estimate of anticipated inflation ($\widehat{\dot{P}^e}$) for that year.[11] From the CPI figures in column 6 one can see that the observation period can be divided into three consecutive periods relative to the level of inflation: 1977-1983 could be considered as a high inflation period, with average inflation at 9.58%, 1984-1992 a medium inflation period, with average inflation at 4.67%, and 1993-1997 a low inflation period, with average inflation at 1.46%. The comparison of the mean (or median) wage-growth figures in columns 3 (or 4) with the CPI figures reveals that there exists a positive relationship between the level of realised inflation and the *location* of the wage-growth distribution across years. Also, the comparison of the standard deviation figures in column 5 with the CPI figures reveals a positive relationship between the level of realised inflation and the *spread* of the wage-growth distribution. We note that DWR leads to a compression of the wage-change distribution. Since DNWR is less likely to hold at higher rates of inflation, this mechanism could explain the positive relationship between the level of realised inflation and the spread of the wage-growth distribution. This positive relation may also arise if expectations about inflation are more diverse during high inflation periods.

Table 2.2 shows the incidence of nominal-wage adjustments relative to the value of zero and the realised level of inflation, by year. Only 102 (or 0.9%) of the contracts in the entire observation period show nominal-wage cuts, while a substantial number (1142 or 10.4%) show a wage freeze; jointly both figures could be considered as strong evidence in favour of the presence of DNWR. The wage freezes are particularly pronounced during the low inflation years; for each of the years 1993-1996 the proportion of contracts with a wage freeze was above 35%, peaking at 51.0% in 1993. On the other hand, 6045 (or 55.2%) of the contracts exhibit negative real-wage growth, while 4801 of them had at the same time positive nominal-wage growth. As expected, the number of contracts that had exactly zero real-wage growth is negligible, just 1 in this case, and the remaining 4899 (or 44.8%) contracts showed both nominal and real-wage increase.

2.4 Empirical Specification

The problem of testing for the presence of a particular type of rigidity using micro data could be stated as one where, having several yearly samples of observations on nominal-wage growth

$$\{\dot{w}_{ti}\}_{\substack{t=1,\ldots,T \\ i=1,\ldots,n_t}} \tag{2.1}$$

where $\dot{w}_{ti}$ represents the nominal-wage growth agreed by the i'th bargaining unit in year t, T is the number of yearly samples and n_t the number of observations in sample t, one would want to test whether these were generated from rigidity-free or rigidity-contaminated

[11] The proxy for anticipated inflation is the one-year-ahead forecast from an AR(6) regression model with a GARCH(1,1) error process.

yearly distributions. Formally, the hypotheses to be tested could be stated as follows

$$\begin{aligned} H_0 &: \; F_t(\dot{w}) = F_t^N(\dot{w}) \\ H_1 &: \; F_t(\dot{w}) = G^R\left(F_t^N(\dot{w})\right) \end{aligned} \tag{2.2}$$

where $F_t(\dot{w})$ is the cdf of the actual-wage-growth distribution, $F_t^N(\dot{w})$ the cdf of the notional-wage-growth distribution, and $G^R\left(F_t^N(\dot{w})\right)$ the cdf of the rigidity-contaminated wage-growth distribution, in year t. The functional $G^R(\cdot)$ is used generically to represent the distortions introduced by the presence of rigidity, which can be either DNWR ($R = n$), or DRWR ($R = r$), or both ($R = nr$).[12]

Exploiting the distinct nature of the distortions in the shape caused by the presence of each type of rigidity, a test for the presence of rigidity of type R could be based on the comparison of the shape of the estimated actual-wage-growth distribution with the shape of the notional distribution (the counterfactual): if there were statistically significant differences in their shape of similar nature to those one would expect to find if rigidity of type R were present, this could be considered as evidence in favour of the presence of rigidity of type R. Formally, this would require one to have information on both $F^N(\cdot)$, that describes the counterfactual distribution, and $G^R(\cdot)$, that characterises the differences due to the presence of rigidity. Obtaining information on the nature of $G^R(\cdot)$ is relatively straightforward, as we have already done informally in Section 2.2. This is not the case for $F^N(\cdot)$, since typically we do not observe the notional-wage growth and thus we cannot make inference about the shape of its distribution *directly*.

The way we proceed here is to use the available actual-wage-growth data to infer information about it *indirectly*, estimating jointly the notional distribution and the distortions due to DWR. The basic idea is to test the hypotheses about the shape of the actual-wage-growth distribution in terms of the heights of the bars of the corresponding probability histogram. Its implementation can be organised in two stages:

Stage 1: Formulation of hypotheses in terms of the parameters of the probability histograms The aim in this stage is to transform the original problem of testing hypotheses about the cdf of the distribution of the actual-wage-growth data from each year in the observation period, as described in (2.2), to one where we test equivalent hypotheses about the corresponding probability histogram.

First we define the probability histograms. Let $P_{jt} \equiv F_t(h_{j+1,t}) - F_t(h_{j,t})$ be the height of the bar of the probability histogram of the actual-wage-growth distribution in year t that corresponds to the j'th bin, denoted by $\mathcal{B}_{jt} \equiv [h_{j,t}, h_{j+1,t}]$, where the bin index $j \in \{-J, \ldots, 0, \ldots, J\}$ indicates the position of the bins in the probability histogram.[13] Given that our analysis aims to examine the shape of the distributions but not their

[12] In this setup, we ignore the presence of measurement error in the wage-growth data. This is a realistic assumption when we work with the Canadian contract data which are collected by the regulating agency HRDC.

[13] Then, the collection of the $2J + 1$ bars defines the probability histogram for that year.

location, j is defined to indicate the position of the bins relative to each other rather than relative to values on the real line. In particular, the bin indexed by $j = 0$ contains the median of the actual-wage-growth distribution, bins indexed by a negative j lie $|j|$ positions to the left of the median bin, and bins indexed by a positive j lie j positions to its right. We refer to the probability histograms defined in this way as 'standardised'.

Having defined the probability histograms in this particular way, in the next step we parameterise P_{jt} under the two hypotheses, i.e.

$$P_{jt} = \begin{cases} p^N\left(z^N_{jt}; b^N_j\right) & \text{, if } H_0 \text{ is true} \\ p^R\left(z^R_{jt}; b^R_j\right) & \text{, if } H_1 \text{ is true} \end{cases} \tag{2.3}$$

where $p^N(\cdot)$ is the function of a vector of observables z^N_{jt} that gives the height of the j'th bar of the probability histogram of the notional distribution in year t, $p^R(\cdot)$ the function of observables z^R_{jt} that gives the height of the corresponding bar of the probability histogram of the rigidity-contaminated distribution in the same year, and b^N_j and b^R_j the corresponding vectors of parameters. Typically both z^N_{jt} and z^R_{jt} will contain dummy variables that indicate the relative position of bin j in the probability histogram,[14] and additional variables that capture characteristics of the year t, while z^R_{jt} will additionally contain variables that indicate the position of bin j relative to the position of the bins containing the values taken by the rigidity bounds in the population.[15]

Given that the shape of the actual-wage-growth distributions is reflected by the height of the bars of the corresponding probability histograms, then, in principle, we could formulate hypotheses about it in terms of the values of the parameters of the functions that describe these heights. Suppose that there is a set of restrictions on the vector of parameters b^R_j, namely $H\left(b^R_j\right) = 0$, such that the two functions $p^N(\cdot)$ and $p^R(\cdot)$ coincide.[16] Then the proposed strategy to test for the presence of rigidity of type R is, firstly, to estimate the parameter vector b^R_j, and then to test the hypotheses

$$\begin{aligned} H_0: &\quad H\left(b^R_j\right) = 0 \\ H_1: &\quad H\left(b^R_j\right) \neq 0 \end{aligned} \tag{2.4}$$

Stage 2: Estimation of the probability histogram parameters and hypothesis testing In this stage we estimate b^R_j and test the hypotheses stated in (2.4). The estimation is done in two steps, and exploits the fact that we have multiple samples on

[14] Therefore, these variables will be functions of j.

[15] Therefore, these variables will be functions of both j and the corresponding indices of the bins that contain the point zero, i.e. the rigidity bound for DNWR, and the anticipated inflation values, i.e. the rigidity bounds for DRWR.

[16] It is natural to think of $G^R(F^N_t(\cdot))$ as the unrestricted case of $F_t(\cdot)$ since $G^R(F^N_t(\cdot)) = F^N_t(\cdot)$ in the special case that $G^R(\cdot)$ is the 'identity' functional. Consequently, we think of of $p^R(z^R_{jt}; b^R_j)$ as the unrestricted model of P_{jt}.

actual-wage-growth.

In Step 1, using the actual-wage-growth data from each year in the observation period, we produce estimates of the heights of the bars of the corresponding probability histograms. Let $\hat{P}_{jt}$ be the estimator of P_{jt} and $\hat{p}_{jt}$ the corresponding estimate. Then this exercise produces a collection of estimates $\{\hat{p}_{jt}\}_{\substack{t=1,\ldots,T \\ j=-J,\ldots,J}}$, which includes an estimate for the height of each bar in the histogram for each year in our sample.

In Step 2, for each j, using the set of T estimates of the height of bar j from all years, i.e. $\{\hat{p}_{jt}\}_{t=1,\ldots,T}$, as the set of 'observations' on $\hat{P}_{jt}$,[17] we estimate the regression of $\hat{P}_{jt}$ on the vector of observables z^R_{jt}. When the estimator $\hat{P}_{jt}$ is unbiased, the regression function will coincide with $p^R\left(z^R_{jt}; b^R_j\right)$, and the regression equation will look like this

$$\hat{P}_{jt} \;=\; E\left(\hat{P}_{jt} \mid z^R_{jt}\right) + \varepsilon_{jt} \;=\; p^R\left(z^R_{jt}; b^R_j\right) + \varepsilon_{jt} \tag{2.5}$$

Therefore, the estimation of this equation would give estimates of the parameter vector b^R_j and its variance-covariance matrix, enabling us to test the restrictions stated in (2.4). In practice the regression equations corresponding to all bar heights are estimated jointly since this is typically more efficient.[18]

For the identification of the parameters of the model it is required that each type of rigidity distorts different parts of the wage-growth distribution at least for some of the years in the sample. In this way, there will be sufficient variation in the dummy variables that indicate the bins that are affected by the distortions, so that these will not be collinear with the dummy variables that indicate the position of the bins in the notional probability histogram.

Our chosen parameterisation for the heights of the bars of the probability histograms under the null hypothesis (i.e. for the notional[19] distribution), is the following

$$\begin{aligned} p^N\left(z^N_{jt}; b^N_j\right) &= \beta_{1|j|} + \beta_{2|j|} \times up_{jt} + \left(\beta_{3|j|} + \beta_{4|j|} \times up_{jt}\right) \times m_t \quad , \quad j \neq 0 \\ &= \beta_{10} + \beta_{30} \times m_t \quad , \quad j = 0 \end{aligned} \tag{2.6}$$

where m_t denotes the median of the actual-wage-growth data in year t, up_{jt} is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ lies to the right of the bin containing the median ($j > 0$), and the β's are coefficients to be estimated. With this parameterisation the $2J+1$ probability bars in each histogram can have different height from each other, therefore the notional distribution is not restricted to have any particular shape, and, in particular, to be symmetric. Furthermore, by making the bar height to be a linear function of the location of the actual-wage-growth distribution, and therefore of the location of the notional

[17] Now $t = 1, \ldots, T$ becomes the observation index.

[18] In such a case, the system would consist of $2J + 1$ equations. The dependent variable corresponding to the equation for a particular observation would be $\hat{P}_{jt}$, where j is the equation index, and t the within equation observation index. To estimate the system we would have in total $(2J+1) \times T$ observations, with T observations on each equation. More details of this case are discussed below.

[19] Given the parameterisation of the probability histograms under the alternative, which we discuss later, here we take the notional distribution to be the the nominal-wage-growth distribution free of any kind of distortions, either due to DWR or menu costs.

distribution itself, we allow for the shape of the notional distribution to vary with its location. For example, suppose that the notional distribution is symmetric around the bin containing m_t and, further, that its spread increases as its centre moves to higher values.[20] Then $\beta_{2|j|}$ and $\beta_{4|j|}$ will be equal to zero due to the symmetry assumption, $\beta_{1|j|}$ will be non-negative, and $\beta_{3|j|}$ will be negative for the bins in the middle of the distribution, i.e. for small $|j|$, and positive for the bins that lie to the tails of the distribution, i.e. for large $|j|$. Alternatively, if we allow $\beta_{4|j|}$ to be non-zero for some values of j, then the skewness of the notional distribution will also vary with the location.[21]

In order to test for the presence of both types of rigidity, the parameterisation of the probability histogram under the alternative hypothesis should reflect the distortions due to the presence of both. We assume that

$$p^R\left(z_{jt}^R; b_j^R\right) = p^N\left(z_{jt}^N; b_j^N\right) + D^u\left(z_{jt}^u; \mu\right) + D^n\left(z_{jt}^n; \gamma\right) + D^r\left(z_{jt}^r; \delta\right) \text{ , for } R = nr \qquad (2.7)$$

where $D^n\left(z_{jt}^n; \gamma\right)$ is defined to be the difference between the height of the j'th bar of the rigidity-contaminated probability histogram and the height of the corresponding bar of the notional probability histogram in year t that is due to the presence of DNWR, and $D^r\left(z_{jt}^r; \delta\right)$ the corresponding difference that is due to the presence of DRWR. We also allow for distortions due to the presence of menu costs, captured by the term $D^u\left(z_{jt}^u; \mu\right)$.

For the effect of DNWR we write

$$D^n\left(z_{jt}^n; \gamma\right) = (\gamma_1 + \gamma_2 \times m_t) \times d0_{jt} + (\gamma_3 + \gamma_4 \times m_t) \times dn_{jt} + \gamma_5 \times dz1_{jt} \qquad (2.8)$$

where $d0_{jt}$ is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ contains the point zero, dn_{jt} a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is to the left of the bin containing the point zero, and $dz1_{jt}$ a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is the first bin to the right of the bin that contains the point zero. With the inclusion of the first term we can capture the distortion that applies to the bin that contains the point of zero nominal-wage-growth, and, with the second term, the distortion that applies to each one of the bins that contain negative values of wage growth.[22] In particular, γ_1 accounts for the distortion associated

[20]This would imply a positive relationship between the spread and location of the histograms of the actual-wage-growth data irrespective of whether any type of rigidity is present or not.

[21]The assumption in the original Kahn (1997) methodology that the shape of the notional distribution is the same across years, has often been cited as one of the main drawbacks of this methodology as in most actual-wage-growth data sets there appears to exist a variation in the spread of the distribution across years characterised by different levels of inflation. This point is raised by Nickell and Quintini (2003) who go on to propose a flexible way of studying DNWR. They exploit the fact that if DNWR is important, the distribution of real-wage change across individual agents should, other things equal, be influenced by inflation and study the proportion of real-wage changes which is below -x%, where x ranges from 2 to 9. This aspect of their analysis focuses on the overall shape of the real-wage-change distribution, rather than the data of the histogram heights used here. Both approaches are very flexible.

[22]Our approach subtracts a constant amount from the height of each bin below zero. Assuming dying out tails, this, as a percentage of mass, becomes larger the further to the left we go implying that large notional cuts are less likely than would be in the case under a proportional specification. Another advantage of this specification is that it is linear in parameters, and therefore easier to estimate, compared to a 'proportional' type of rigidity.

with the bin that contains the point of zero nominal-wage-growth and γ_3 the distortion associated with the bins that lie to the left of this bin in the special case where the centre of the notional distribution, which we proxy by m_t, is located at the point zero (i.e. $m_t = 0$). In that case, and, in the presence of DNWR, we would expect γ_1 to be positive, signifying the concentration of probability mass surplus in the zero nominal-wage-growth bin, and γ_3 negative, signifying the loss of probability mass from the bins that contain negative values of notional-wage growth. When the centre of the notional distribution is located further to the right ($m_t > 0$), a smaller part of the left tail of the notional distribution lies below zero, i.e. the proportion of notional-wage cuts falls, and, therefore the proportion of notional-wage changes that become wage freezes due to DNWR is expected to fall. In that case, γ_2 must be negative, signifying the reduction in the probability mass surplus in the zero-nominal-wage-growth bin, while γ_4 could be either positive or negative or zero, as the amount of mass deficit from each bin containing negative values could change in any direction relative to its level at $m_t = 0$. The inclusion of the last term enables us to test the hypothesis that, apart from shifting mass to the point of zero nominal-wage-growth, the presence of DNWR could also induce a shift of mass beyond the point zero, towards small positive values (in that case, $\gamma_5 > 0$).[23]

The distortion in the height of the probability bar of bin $\mathcal{B}_{jt}$ due to DRWR is assumed to be given by

$$D^r\left(z^r_{jt};\delta\right) = \begin{cases} \delta_{1k} + \delta_{2k} \times J^P_t, \quad k = j - J^P_t & , \text{ if } k_{\min} \leq k \leq k_{\max} \\ 0 & , \text{ otherwise} \end{cases} \tag{2.9}$$

where J^P_t is the value of the index of the bin in year t that contains the centre of the anticipated-inflation distribution in that year[24] and, thus, k the distance between bin $\mathcal{B}_{jt}$ and that bin.[25] It is more convenient to write (2.9) more compactly as follows

$$D^r\left(z^r_{jt};\delta\right) = \sum_{\nu=k_{\min}}^{k_{\max}} \left(\delta_{1\nu} + \delta_{2\nu} \times J^P_t\right) \times dp_{\nu,jt} \tag{2.10}$$

where $dp_{\nu,jt}$ are dummy variables indicating whether bin $\mathcal{B}_{jt}$ is located k positions from

[23] Holden (1989,2004) and Cramton and Tracy (1992) describe mechanisms of DNWR where unions can not only resist nominal-wage cuts but also induce small positive nominal-wage changes by threatening to work less efficiently during bargaining (holdout). Cramton and Tracy (1992) find empirical support for this model in US wage contract data. Holden (1989,1998) do the same for wage setting in the Nordic countries.

[24] In the empirical application we have proxied this value either with the realised inflation in year t, measured by the CPI in year t, or a GARCH estimate of anticipated inflation in year t. See Table 2.1 for their values.

[25] The index k is assumed to take values from the set $\{k_{\min}, \ldots, 0, \ldots, k_{\max}\}$. The bin for which $k = 0$ contains the centre of the anticipated-inflation distribution, bins with positive values of k are located to the right of this bin, and bins with negative values to its left. The values taken by k_{min} and k_{max} are determined empirically.

the bin that contains the centre of the anticipated-inflation distribution in year t,

$$dp_{\nu,jt} = \begin{cases} 1 & \text{if } \nu = k \left(= j - J_t^P\right) \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

With this specification we allow for the size of the distortions to differ according to the location of the bin in the support of the anticipated-inflation distribution (through the indexing by k), and its location in the support of the notional-wage-growth distribution (through the dependence on J_t^P). In the presence of DRWR, the δ_{1k}'s, which account for the distortion when the centre of the anticipated-inflation distribution is located in the same bin as the median of the actual-wage-growth distribution ($J_t^P = 0$), are expected to be positive for the largest (and positive) values of k and negative for the smallest (and negative) values of k, signifying the shift of probability mass towards the right end of the support of the anticipated-inflation distribution - see discussion on p. 14. When J_t^P takes different values, the values of the δ_{2k}'s must be such that the distortions $(\delta_{1k} + \delta_{2k} \times J_t^P)$ are qualitatively similar to the case where $J_t^P = 0$, however no specific statements can be made about their sign or size unless specific assumptions are made about the nature of the joint distribution of the notional-wage growth and anticipated inflation, and the rigidity mechanism.

Finally the effect of menu costs is parameterised as follows

$$D^u\left(z_{jt}^u; \mu\right) = \mu \times dnp1_{jt} \tag{2.12}$$

where $dnp1_{jt}$ is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is either one position to the left or to the right of the bin that contains the point zero. Therefore we allow for a symmetric loss of mass ($\mu < 0$) around and close to zero.[26]

2.5 Estimation

To produce the estimates of the heights of the bars of the probability histograms ($\hat{p}_{jt}$), we use the proportion of observations in the sample for year t that fall in bin j as the estimator

[26] Christofides and Leung (2003) found only weak evidence for the presence of menu costs for this data. However, accounting for their presence in our specification ensures that the magnitude of parameter γ_5, which measures the distortion in the first bin to the right of the bin that contains the point zero, attributed to DNWR, is estimated correctly. Given the model setup, we are not able test formally for the concentration of mass around the point zero within the bin that contains it. However data inspection shows that almost all mass accumulated in that bin is attributed to point zero.

of P_{jt}.[27] This estimator, denoted by $\hat{P}_{jt}$, could be defined as

$$\hat{P}_{jt} \equiv \sum_{i=1}^{n_t} \frac{d_{jti}}{n_t} \tag{2.13}$$

where d_{jti} is a dummy variable that takes the value of 1 if $\dot{w}_{ti}$ falls in bin $\mathcal{B}_{jt}$ and 0 otherwise, and n_t is the number of observations in year t. Since $\Pr(d_{jti} = 1) = \Pr(\dot{w}_{ti} \in \mathcal{B}_{jt}) = P_{jt}$, then d_{jti} is a Bernoulli random variable with mean P_{jt}. Furthermore, assuming that $\dot{w}_{ti} \overset{iid}{\sim} F_t(\dot{w})$ within t, $\hat{P}_{jt}$ becomes the sample mean of i.i.d. *Bernoulli*(P_{jt}) random variables and is thus an unbiased estimator of the true height P_{jt}, as well as consistent and asymptotically normal. We can further derive the exact algebraic expression for the covariance between any pair of estimators that correspond to bins from the same or different probability histograms. Treating the wage growth associated with different bargaining pairs as being independent also across years, this expression takes the form

$$Cov\left(\hat{P}_{jt}, \hat{P}_{\zeta\tau}\right) = \begin{cases} \frac{P_{jt}(1-P_{jt})}{n_t} & , t = \tau \text{ and } j = \zeta \\ -\frac{P_{jt}P_{\zeta t}}{n_t} & , t = \tau \text{ and } j \neq \zeta \\ \sum_{i \in \mathcal{I}_t \cap \mathcal{I}_\tau} \frac{\Pr\left(d_{jti} = d_{\zeta\tau i} = 1\right) - P_{jt}P_{\zeta\tau}}{n_t n_\tau} & , t \neq \tau \end{cases} \tag{2.14}$$

where $\mathcal{I}_t$ and $\mathcal{I}_\tau$ are the sets of indices denoting the bargaining pairs which appear in our sample to have a contract agreement in years t and τ respectively,[28] while $j, \zeta \in \{-J, \ldots, J\}$ and $t, \tau \in \{1, \ldots, T\}$.[29]

In Step 2 of the estimation stage we treat the $2J + 1$ equations as a system. After imposing the cross-equation parameter restrictions implied by the parameterisation of (2.7),[30] the equation for a typical observation for the stacked data can be written as follows

$$\hat{P}_{jt} = \sum_{q=-J}^{J} p^N\left(z_{qt}^N; b_q^N\right) \times d_{qt}^* + D^u\left(z_{jt}^u; \mu\right) + D^n\left(z_{jt}^n; \gamma\right) + D^r\left(z_{jt}^r; \delta\right) + \varepsilon_{jt} \tag{2.15}$$

where d_{qt}^* is a dummy variable that is equal to 1 if $q = j$, and 0 otherwise. In matrix form

[27] We use the median of the actual-wage-growth data from year t, denoted by $\hat{m}_t$, as an estimate of m_t. Therefore the bin of the estimated probability histogram indexed by $j = 0$ is the one that contains $\hat{m}_t$. The bin width is set to be equal to 1% and the bin endpoints take values from the set $\{\ldots, -1.5, -0.5, 0.5, 1.5, \ldots\}$. Thus, the point zero is at the centre of the bin that contains it and small wage changes around zero also fall in the same bin. Furthermore, J is chosen to be equal to 8 so that each probability histogram consists of 17 bins that cover in total an interval of 17 percentage points. In this way we achieve coverage of more than 97% of the data points in each yearly sample.

[28] With this notation we allow for the presence of agreements for the same bargaining unit in several of the available yearly samples.

[29] Putting some intuition in this result, first we note that the estimators are sample means of non-identically distributed Bernoulli variables for samples of different sizes and, therefore, should be expected to have different variances, as shown in the first line. Also, since the heights of the bars for each histogram must sum up to one, then the estimators of these heights should be expected to be negatively correlated, as they appear to be in the second line. Finally, to the extent that the wage settlements reached by the same bargaining unit at different points in time are correlated with each other, then the probability estimators in different years could also be correlated, as suggested by the result in the third line.

[30] Specifically, the parameter vectors that capture the effect of the rigidities, i.e. γ and δ, and the parameter that captures the effect of menu costs, i.e. the μ, are common to all equations.

the system can be written as

$$\hat{\mathbf{P}} = \mathbf{Z}\mathbf{b} + \varepsilon \tag{2.16}$$

where $\hat{\mathbf{P}} \equiv \begin{bmatrix} \hat{\mathbf{P}}_{-J} & \hat{\mathbf{P}}_{-J+1} & \cdots & \hat{\mathbf{P}}_{0} & \cdots & \hat{\mathbf{P}}_{J-1} & \hat{\mathbf{P}}_{J} \end{bmatrix}'$ is the vector of dependent variables for the entire system, and $\hat{\mathbf{P}}_j \equiv \begin{bmatrix} \hat{P}_{j1} & \hat{P}_{j2} & \cdots & \hat{P}_{jT} \end{bmatrix}$ the vector of dependent variables that corresponds to equation j.

The choice of optimal estimation method for the parameters of the system depends on the nature of the variance-covariance matrix of the vector $\hat{\mathbf{P}}$ of estimators, denoted by $Var\left(\hat{\mathbf{P}}\right)$, whose typical element is $Cov\left(\hat{P}_{jt}, \hat{P}_{\zeta\tau}\right)$. Clearly from (2.14) we see that $Var\left(\hat{\mathbf{P}}\right)$ is not spherical, since the diagonal elements are not identical (first line of the result), and since there are also some non-zero off-diagonal elements (second and third lines). Therefore the Ordinary Least Squares (OLS) procedure, despite producing consistent estimates of $\mathbf{b}$, gives wrong standard error estimates. Therefore we opt for the Feasible Generalised Least Squares (FGLS) procedure, substituting the probabilities in the right-hand side of (2.14) with consistent estimates; in the case of the probabilities of the form P_{jt} with the estimates obtained in stage 2 (i.e. $\hat{p}_{jt}$), and for $\Pr\left(d_{jti} = d_{\zeta\tau i} = 1\right)$ with estimates produced in a similar way

$$\Pr\left(\widehat{d_{jti} = d_{\zeta\tau i}} = 1\right) = \sum_{i \in \mathcal{I}_t \cap \mathcal{I}_\tau} \frac{d_{jti} d_{\zeta\tau i}}{\#\left(\mathcal{I}_t \cap \mathcal{I}_\tau\right)} \xrightarrow{p} \Pr\left(d_{jti} = d_{\zeta\tau i} = 1\right) \tag{2.17}$$

where $\#\left(\mathcal{I}_t \cap \mathcal{I}_\tau\right)$ is the number of elements in the set $\mathcal{I}_t \cap \mathcal{I}_\tau$.[31]

2.6 Results

In Table 2.3 we present the estimation results when we apply the FGLS estimator (columns 2 and 3), and the OLS estimator with corrected (columns 4 and 5) and uncorrected standard errors (columns 6 and 7). To obtain these results we have used the GARCH approach, rather than the actual CPI growth, to estimate the mean anticipated inflation rate.[32] The table is divided in three panels; the top panel includes the estimates associated with the notional distribution (β's) and the distortion due to menu costs (μ), the middle panel those associated with the distortion due to the presence of DNWR (γ's), and the bottom panel those associated with the distortion due to the presence of DRWR (δ's). Furthermore, in Table 2.4, we present the results from testing joint hypotheses about the parameters of the model using the Wald and F statistics, which are based on the results from the FGLS estimation. Next we discuss the results from the FGLS estimation.[33]

[31] In order to obtain the results described in the next section, we have assumed - a priori - that all these quantities are equal to zero. Given that our expectation is that the corresponding estimates would have been relatively small and close to zero, we do not think that this assumption has affected considerably these results.

[32] This is more consistent with the conceptual approach of Section 2.2 and produces clearer results.

[33] The OLS results with corrected standard errors are, at least qualitatively, similar; they show a shift of probability mass to the right towards the values of anticipated inflation, a spike at zero, and mass deficit

Notional distribution & menu costs Overall, the majority of the estimates of the β's in the top panel of Table 2.3 are statistically significant, suggesting that the shape of the notional distribution is not fixed but varies with its location.

In particular, the estimates of $\{\beta_{10}, \ldots, \beta_{18}\}$ and $\{\beta_{21}, \ldots, \beta_{28}\}$ suggest that, if the median of the actual-wage-growth distribution were equal to zero, then the corresponding notional distribution would be bell shaped, with the mode of the distribution being located close to the median of the actual distribution. Also, the negative sign of $\{\beta_{21}, \ldots, \beta_{28}\}$ suggests that the distribution has less mass above the bin that contains the actual-wage-growth median.[34] As the median of the actual distribution increases, and therefore the location of the notional distribution also moves to the right, the estimates of $\{\beta_{30}, \ldots, \beta_{38}\}$ and $\{\beta_{41}, \ldots, \beta_{48}\}$ together suggest that there is a flattening of the shape of the distribution in the centre and that there is progressively more concentration of probability mass at the right tail of the notional distribution and less mass at the left tail. The overall effect is that, as the location of the notional distribution moves further to the right, the distribution becomes progressively skewed to the right and its spread increases.

The estimate of μ is statistically significant but suggests that the number of small positive and negative wage adjustments that do not take place, presumably due to menu costs, is relatively small. Specifically, the number of adjustments in the intervals $[-1.5, -0.5]$ and $[0.5, 1.5]$ that do not take place corresponds to around 1.5% of the total number of contracts in each year, confirming an earlier result by Christofides and Leung (2003) for modest menu costs.

DNWR The estimates of the parameters measuring the effect due to the presence of DNWR in the middle panel are also significant and have signs that are consistent with the presence of this type of rigidity; That is, $\hat{\gamma}_1$ is positive, $\hat{\gamma}_2$ negative, and $\hat{\gamma}_3$ negative.

If the median of the actual-wage-change distribution were equal to zero, the bin containing the point of zero nominal-wage growth would attract an estimated excess probability mass of 10.43% ($= \hat{\gamma}_1$), while each bin that contained negative values of wage growth would show a probability adjustment of -1.74% ($= \hat{\gamma}_3$). For all other values of the actual median, the surplus concentrated at the bin of zero nominal-wage growth would change by -1.64% ($= \hat{\gamma}_2$) for each 1% increase in the actual median, while the deficit in the bins containing negative values would become smaller by 0.13% ($= \hat{\gamma}_4$). These results are similar to those reported by Christofides and Leung (2003) who apply a variant of the original Kahn (1997) approach to test for the presence of DNWR only. The estimate of γ_5 is positive, but small and statistically insignificant. Therefore we only find weak support for the hypothesis that agents can induce small positive nominal-wage changes, in the fashion described by Holden (1989,2004) and Cramton and Tracy (1992).

below zero. For the OLS results without correction of the standard errors we note that all of the parameters that measure the effect of DRWR, with the exception of one, are statistically insignificant.

[34] If we believed that the median of the notional distribution were close enough to the median of the actual so that they were both located in the same bin, then we could interpret this result as one that suggests that the notional distribution is non-symmetric.

DRWR The majority of estimates of the parameters in the bottom panel of Table 2.3 are statistically significant, suggesting that the shape of the wage-growth distribution is distorted in the region where we expect future inflation beliefs to lie. Furthermore, the pattern of the distortions suggests a shift of probability mass to the right towards these values, that is similar to what we would expect if DRWR were present.

In particular, when the centre of the anticipated-inflation distribution is located in the same bin as the median of the actual distribution, then this bin ($j = 0$ & $k = 0$), as well as the bin that lies immediately to its left ($k = -1$) and all the bins to its right ($k = 1, \dots, 5$) attract a surplus of probability mass that is statistically significant. At the same time, the remaining bins that lie to the left of this group of bins ($k = -2, \dots, -5$) show a deficit of probability mass, that is also statistically significant in all cases except for the leftmost bin ($k = -5$). The bin indexed by $k = -1$ attracts the biggest surplus, equal to 8.71%, while the bin at the centre of both the actual and the anticipated-inflation distributions attracts the second biggest, equal to 7.78%. The surplus continues to diminish as we move further to the right, taking values between 5.44% to 0.76%. On the other hand, the deficit for the bins indexed by $k = -2, \dots, -4$, ranges between 0.99% and 1.61%.

This pattern remains qualitatively unchanged when the centre of the distribution of anticipated inflation is located at its sample values,[35] while quantitatively the variation is relatively very small. The signs of the estimates of $\{\delta_{-25}, \dots, \delta_{25}\}$ suggest that quantitatively the distortions become more pronounced at the centre of the anticipated-inflation distribution the further to the right the centre of this distribution is located relative to the centre of the actual-wage-growth distribution. In particular, the excess mass becomes progressively more concentrated in the two bins at its centre, indexed by $k = 0, -1$, while the deficit in the bins further to their left increases.

Joint tests The test results in Table 2.4 concerning joint hypotheses about subsets of the parameter set offer additional support to our main conclusions about the presence of the two types of rigidity and the shape of the notional distribution. In particular, we reject overwhelmingly the null hypotheses: (i) that the shape of the actual-wage-growth distribution coincides with the shape of the rigidity-free distribution (line 1), (ii) the absence of DRWR (line 2), and (iii) the absence of DNWR (line 3). We also reject the hypotheses: (iv) that the shape of the notional distribution remains the same as the centre of the distribution changes location (line 4), and (v) that the notional distribution is symmetric around the bin containing the median of the actual distribution (line 5).

Because of the large number or estimated parameters and the presence of interaction terms, the nature of the estimated distortions due to the presence of both types of rigidity may not be immediately clear. Therefore, in Figure 2.2, we draw the fitted probability histograms for the notional and actual-wage-growth distributions, for selected years of high, medium, and low inflation. The histograms in this figure are 'standardised', therefore the

[35] These appear in the last column of Table 2.1.

bin indexed by zero contains the median of the actual-wage-growth data from the relevant year. The light-shaded bars correspond to the notional distribution, i.e. $p^N\left(z_{jt}^N;\hat{b}_j^N\right)$, while the dark-shaded bars to the actual distribution, i.e. $p^R\left(z_{jt}^R;\hat{b}_j^R\right)$. In the diagrams we also indicate (except for the top row when it is off the left point of the chart) which bins contain the point zero ("0") and the estimated centre of the distribution of anticipated inflation ("$\widehat{\dot{P}_t^e}$"). These diagrams clearly demonstrate a shift of probability mass to the right, towards the values of anticipated inflation in the population, that produces a similar type of distortions to those in the simulated diagrams that assume the presence of DRWR, in Figure 2.1. Furthermore, we can distinguish a spike at the bin containing the point zero, for example bin -5 in the graphs for years 1983 and 1989, bin -4 for 1984, and bin -2 for 1992, and deficit in the bins below zero, which is consistent with the presence of DNWR.[36]

2.7 Conclusion

In this paper, we study collective bargaining wage outcomes drawn from the Canadian unionised sector, over a long period of diverse inflation experience. Earlier studies involving this data found evidence for the presence of DNWR. The challenge was to specify mechanisms consistent with the notion of DRWR and to superimpose these mechanisms on the broad approach used to measure DNWR in the past. The results obtained suggest that DRWR is clearly present in the data and that it can be identified over and above substantial DNWR effects.

[36]We have estimated several variations of the model in order to examine the sensitivity of our results to the particular specification of the model and found that these were robust, at least in a qualitative sense; in all cases we estimated a shift of mass to the right towards point zero and the interval we expect to contain the values of anticipated inflation, the presence of excess mass at zero, and a mass deficit in the part of the distribution that lies below zero.

Year	#	WNC+COLA			CPI	$\widehat{\dot{P}^e}$
		mean	*median*	*s.d.*		
1977	226	8.69	8.20	2.99	7.55	7.22
1978	673	8.16	7.43	2.89	8.01	8.42
1979	569	10.64	10.11	3.18	8.95	8.45
1980	520	12.39	11.95	2.87	9.13	9.28
1981	450	13.64	13.10	3.20	10.16	11.66
1982	562	10.31	10.69	3.64	12.43	10.43
1983	643	4.89	5.00	2.63	10.80	6.05
1984	676	3.76	4.00	1.90	5.86	4.50
1985	519	3.78	4.04	2.14	4.30	3.81
1986	551	3.65	4.10	1.82	3.96	4.08
1987	557	3.90	3.83	1.77	4.18	4.37
1988	556	4.92	4.89	1.72	4.34	3.97
1989	493	5.68	5.22	1.82	4.05	4.83
1990	547	5.79	5.77	1.99	4.99	4.55
1991	530	3.89	4.19	2.19	4.76	5.91
1992	632	2.16	2.00	1.80	5.62	1.49
1993	516	0.75	0.00	1.35	1.49	2.00
1994	471	0.60	0.00	1.63	1.86	0.50
1995	460	0.86	0.68	1.16	0.16	2.24
1996	448	1.22	0.87	1.31	2.16	1.43
1997	346	1.87	1.87	1.35	1.62	1.95
Total	10945					

Table 2.1: Descriptive statistics.

Year	$\dot{w} < 0$		$\dot{w} = 0$		$0 < \dot{w} < CPI$		$\dot{w} = CPI$		$\dot{w} > CPI$	
	#	%	#	%	#	%	#	%	#	%
1977			2	*0.9*	86	*38.1*			138	*61.1*
1978					393	*58.4*			280	*41.6*
1979					198	*34.8*			371	*65.2*
1980					43	*8.3*			477	*91.7*
1981			1	*0.2*	38	*8.4*			411	*91.3*
1982	1	*0.2*	3	*0.5*	397	*70.6*			161	*28.6*
1983	4	*0.6*	26	*4.0*	597	*92.8*			16	*2.5*
1984	1	*0.1*	61	*9.0*	559	*82.7*			55	*8.1*
1985	1	*0.2*	26	*5.0*	286	*55.1*			206	*39.7*
1986	2	*0.4*	24	*4.4*	238	*43.2*	1	*0.2*	286	*51.9*
1987			17	*3.1*	307	*55.1*			233	*41.8*
1988			4	*0.7*	203	*36.5*			349	*62.8*
1989					60	*12.2*			433	*87.8*
1990			14	*2.6*	136	*24.9*			397	*72.6*
1991	2	*0.4*	57	*10.8*	243	*45.8*			228	*43.0*
1992	7	*1.1*	82	*13.0*	488	*77.2*			55	*8.7*
1993	18	*3.5*	263	*51.0*	116	*22.5*			119	*23.1*
1994	53	*11.3*	186	*39.5*	146	*31.0*			86	*18.3*
1995	9	*2.0*	162	*35.2*	2	*0.4*			287	*62.4*
1996	3	*0.7*	164	*36.6*	174	*38.8*			107	*23.9*
1997	1	*0.3*	50	*14.5*	91	*26.3*			204	*59.0*
Total	102	*0.9*	1142	*10.4*	4801	*43.9*	1	*0.0*	4899	*44.8*

Table 2.2: Wage-growth statistics.

	FGLS		OLS-corrected		OLS	
Parameter	**Estimate**	**(Std. Err.)**	**Estimate**	**(Std. Err.)**	**Estimate**	**(Std. Err.)**
β_{10}	0.3112**	(0.0098)	0.2998**	(0.0128)	0.2998**	(0.0293)
β_{11}	0.0881**	(0.0057)	0.2047**	(0.0102)	0.2047**	(0.0282)
β_{12}	0.0418**	(0.0038)	0.1028**	(0.0077)	0.1028**	(0.0258)
β_{13}	0.0385**	(0.0032)	0.0875**	(0.0059)	0.0875**	(0.0233)
β_{14}	0.0250**	(0.0021)	0.0564**	(0.0046)	0.0564**	(0.0204)
β_{15}	0.0194**	(0.0021)	0.0520**	(0.0041)	0.0520**	(0.0195)
β_{16}	0.0155**	(0.0053)	0.0492**	(0.0040)	0.0492*	(0.0200)
β_{17}	0.0195**	(0.0028)	0.0492**	(0.0039)	0.0492*	(0.0204)
β_{18}	0.0183**	(0.0030)	0.0494**	(0.0039)	0.0494*	(0.0205)
β_{21}	0.1105**	(0.0101)	0.0012	(0.0124)	0.0012	(0.0228)
β_{22}	-0.0086	(0.0078)	-0.0273**	(0.0103)	-0.0273	(0.0278)
β_{23}	-0.0448**	(0.0058)	-0.0435**	(0.0084)	-0.0435	(0.0295)
β_{24}	-0.0369**	(0.0038)	-0.0382**	(0.0066)	-0.0382	(0.0277)
β_{25}	-0.0267**	(0.0031)	-0.0492**	(0.0050)	-0.0492†	(0.0252)
β_{26}	-0.0187**	(0.0056)	-0.0541**	(0.0042)	-0.0541*	(0.0236)
β_{27}	-0.0174**	(0.0030)	-0.0513**	(0.0041)	-0.0513*	(0.0236)
β_{28}	-0.0185**	(0.0036)	-0.0504**	(0.0040)	-0.0504*	(0.0235)
β_{30}	-0.0204**	(0.0011)	-0.0140**	(0.0012)	-0.0140**	(0.0022)
β_{31}	0.0027**	(0.0008)	-0.0060**	(0.0011)	-0.0060**	(0.0023)
β_{32}	0.0064**	(0.0006)	0.0024**	(0.0009)	0.0024	(0.0024)
β_{33}	0.0018**	(0.0005)	-0.0011†	(0.0007)	-0.0011	(0.0025)
β_{34}	0.0002	(0.0004)	-0.0012*	(0.0006)	-0.0012	(0.0026)
β_{35}	-0.0018**	(0.0003)	-0.0016**	(0.0005)	-0.0016	(0.0025)
β_{36}	-0.0009	(0.0006)	-0.0029**	(0.0004)	-0.0029	(0.0025)
β_{37}	-0.0013**	(0.0003)	-0.0033**	(0.0003)	-0.0033	(0.0023)
β_{38}	-0.0012**	(0.0004)	-0.0034**	(0.0003)	-0.0034	(0.0023)
β_{41}	-0.0113**	(0.0014)	0.0003	(0.0017)	0.0003	(0.0030)
β_{42}	-0.0012	(0.0011)	0.0019	(0.0014)	0.0019	(0.0033)
β_{43}	0.0024**	(0.0008)	0.0026**	(0.0010)	0.0026	(0.0033)
β_{44}	0.0037**	(0.0006)	0.0036**	(0.0008)	0.0036	(0.0033)
β_{45}	0.0044**	(0.0005)	0.0038**	(0.0007)	0.0038	(0.0033)
β_{46}	0.0027**	(0.0007)	0.0055**	(0.0005)	0.0055†	(0.0031)
β_{47}	0.0017**	(0.0003)	0.0049**	(0.0004)	0.0049†	(0.0029)
β_{48}	0.0020**	(0.0005)	0.0042**	(0.0004)	0.0042	(0.0029)
μ	-0.0077**	(0.0022)	-0.0196**	(0.0016)	-0.0196†	(0.0115)
γ_1	0.1043**	(0.0053)	0.1735**	(0.0104)	0.1735**	(0.0176)
γ_2	-0.0164**	(0.0010)	-0.0309**	(0.0019)	-0.0309**	(0.0037)
γ_3	-0.0174**	(0.0016)	-0.0530**	(0.0038)	-0.0530**	(0.0154)
γ_4	0.0013**	(0.0003)	0.0057**	(0.0003)	0.0057*	(0.0023)
γ_5	0.0027	(0.0030)	0.0041	(0.0049)	0.0041	(0.0161)
δ_{-15}	-0.0025	(0.0016)	-0.0084**	(0.0022)	-0.0084	(0.0111)
δ_{-14}	-0.0099**	(0.0020)	-0.0236**	(0.0040)	-0.0236	(0.0151)
δ_{-13}	-0.0161**	(0.0030)	-0.0306**	(0.0059)	-0.0306	(0.0199)
δ_{-12}	-0.0118**	(0.0045)	-0.0323**	(0.0075)	-0.0323	(0.0235)
δ_{-11}	0.0871**	(0.0067)	0.0120	(0.0094)	0.0120	(0.0259)
δ_{10}	0.0778**	(0.0075)	0.0260**	(0.0100)	0.0260	(0.0265)
δ_{11}	0.0544**	(0.0067)	0.0111	(0.0092)	0.0111	(0.0254)
δ_{12}	0.0416**	(0.0058)	-0.0014	(0.0079)	-0.0014	(0.0228)
δ_{13}	0.0199**	(0.0044)	-0.0124*	(0.0062)	-0.0124	(0.0189)
δ_{14}	0.0114**	(0.0029)	-0.0094*	(0.0042)	-0.0094	(0.0142)
δ_{15}	0.0076**	(0.0016)	-0.0007	(0.0025)	-0.0007	(0.0106)
δ_{-25}	-0.0007	(0.0011)	-0.0039*	(0.0016)	-0.0039	(0.0079)
δ_{-24}	-0.0035**	(0.0012)	-0.0051*	(0.0021)	-0.0051	(0.0085)
δ_{-23}	-0.0094**	(0.0015)	-0.0094**	(0.0025)	-0.0094	(0.0087)
δ_{-22}	-0.0043*	(0.0018)	-0.0013	(0.0028)	-0.0013	(0.0085)
δ_{-21}	0.0302**	(0.0027)	0.0149**	(0.0036)	0.0149†	(0.0082)
δ_{20}	0.0099**	(0.0032)	-0.0004	(0.0036)	-0.0004	(0.0080)
δ_{21}	-0.0016	(0.0028)	-0.0009	(0.0034)	-0.0009	(0.0081)
δ_{22}	-0.0062*	(0.0025)	-0.0009	(0.0033)	-0.0009	(0.0084)
δ_{23}	-0.0056**	(0.0021)	0.0013	(0.0031)	0.0013	(0.0086)
δ_{24}	-0.0037*	(0.0016)	0.0091**	(0.0026)	0.0091	(0.0084)
δ_{25}	-0.0018	(0.0013)	0.0018	(0.0022)	0.0018	(0.0077)
N	357		357		357	

Significance levels : † : 10% $*$: 5% $**$: 1%

Table 2.3: Estimation results.

#	W	$Pr(\chi^2_q > W)$	F	$Pr(F_{(q,n-k)} > F)$
1^a	1484.7550	0	54.99091	0
2^b	649.3472	0	29.51578	0
3^c	587.1228	0	117.42460	0
4^d	781.0962	0	45.94683	0
5^e	439.8713	0	27.49195	0

$^a H_0 : \gamma = 0 \cap \delta = 0, H_1 : \gamma \neq 0 \cup \delta \neq 0 \quad (q = 27,\ n - k = 296)$
$^b H_0 : \delta = 0, H_1 : \delta \neq 0 \quad (q = 22,\ n - k = 296)$
$^c H_0 : \gamma = 0, H_1 : \gamma \neq 0 \quad (q = 5,\ n - k = 296)$
$^d H_0 : \beta_{3|j|} = \beta_{4|j|} = 0, H_1 : \beta_{3|j|} \neq 0 \cup \beta_{4|j|} \neq 0,\ \forall j \quad (q = 17,\ n - k = 296)$
$^e H_0 : \beta_{2|j|} = \beta_{4|j|} = 0, H_1 : \beta_{2|j|} \neq 0 \cup \beta_{4|j|} \neq 0,\ \forall j \quad (q = 16,\ n - k = 296)$

Table 2.4: Joint test results.

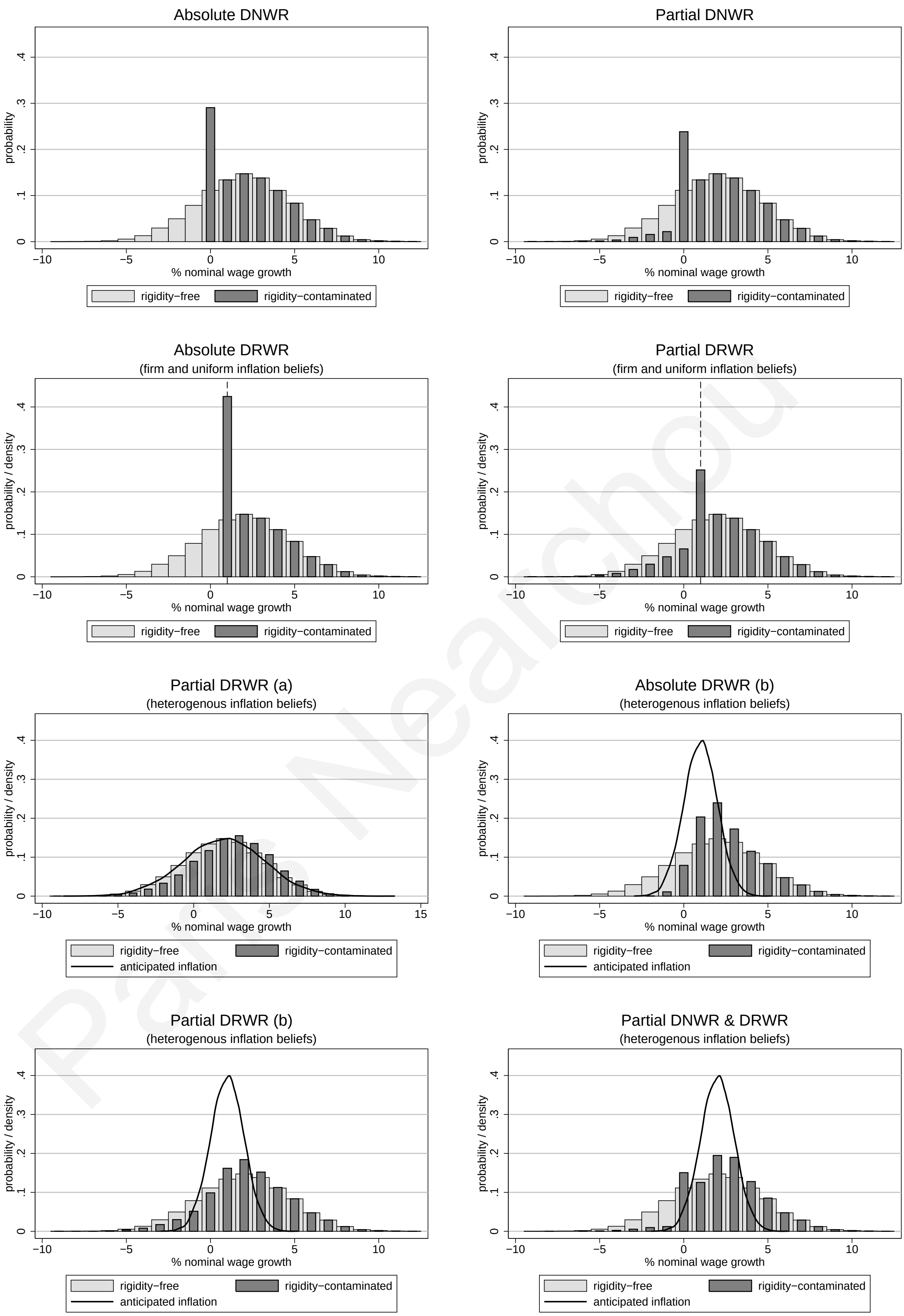


Figure 2.1: Simulated rigidity-free (notional) and rigidity-contaminated nominal WGDs.

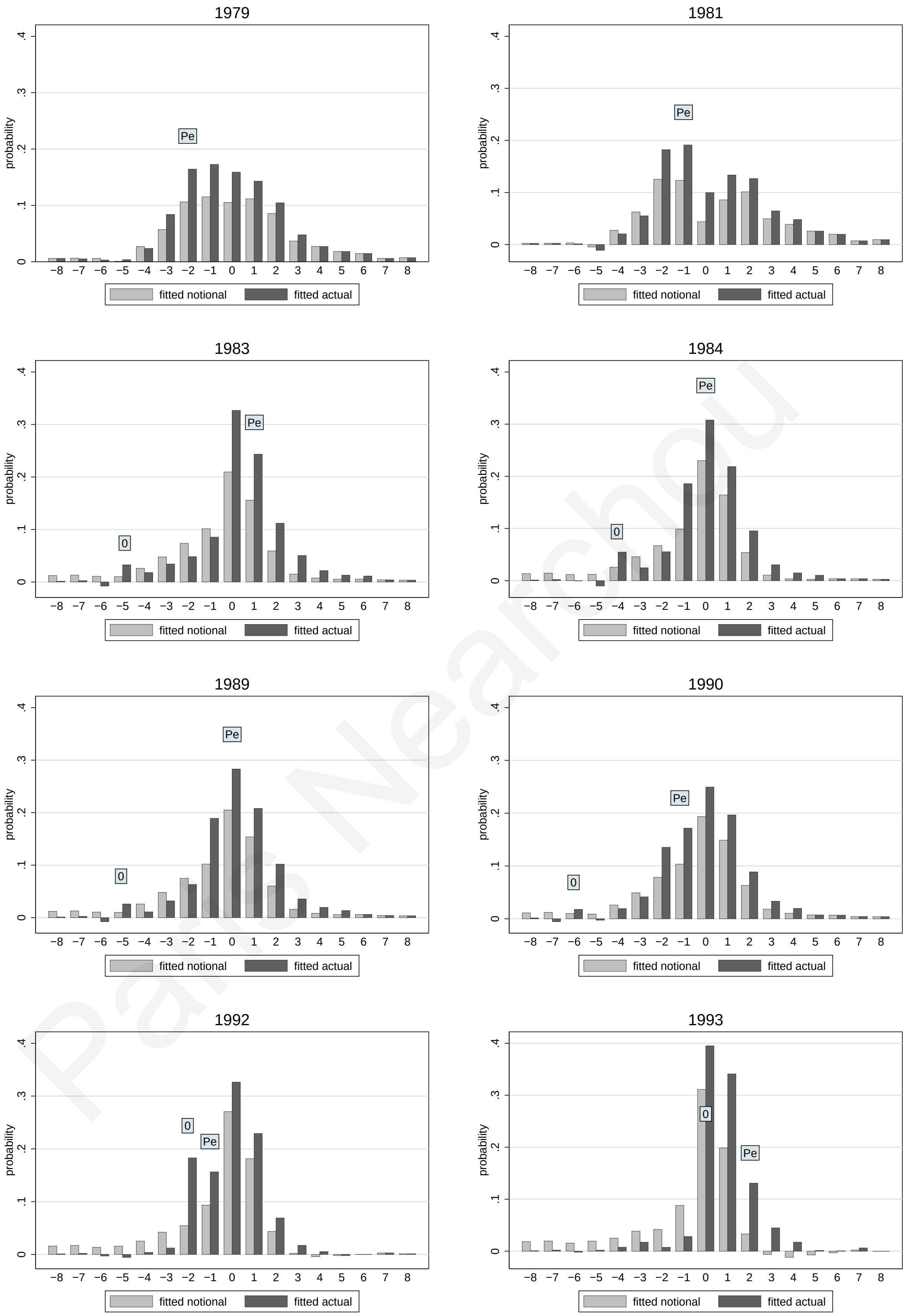


Figure 2.2: Notional Vs Actual nominal-wage-growth distributions (fitted values).

Chapter 3

Patterns of Nominal and Real Wage Rigidity

3.1 Introduction

In Keynesian models, the notion of downward nominal wage rigidity (DNWR) plays an important role in 'rationalising' the failure of the labour market to clear and the existence of unemployment. In these models, employment is determined along the labour demand curve (the 'short' side of the market) and, since increases in employment require movements along this demand curve, the real wage behaves counter-cyclically. Early attempts (Dunlop (1938) and Tarshis (1939)), but also more recent papers (Solon, Barsky and Parker (1994) and Abraham and Haltiwanger (1995), examine nominal wage rigidity indirectly by looking at the cyclical properties of the real wage rate.[1]

McLaughlin's (1994) paper shifted attention to wage growth distributions (WGD) for individuals in panel data, thus giving rise to a more inductive approach to this issue: What do data on the earnings of individuals over time imply about wage rigidity? The papers by, *inter alia*, Lebow, Stockton and Wascher (1995), Fortin (1996), Kahn (1997), Card and Hyslop (1997), Crawford and Harrison (1998), Smith (2000), Altonji and Devereux (2000), Christofides and Stengos (2001, 2002, 2003), Christofides and Leung (2003), Christofides and Li (2005), Dickens and Groshen (2004), and Holden and Wulfsberg (2007) all follow this broad approach. Some of these papers deal not only with DNWR but, also, with aspects of real wage rigidity.

The extent to which DNWR and downward real wage rigidity (DRWR) coexist and interact are points that are worth investigating further.[2] Papers which address both types of rigidity based on the maximum likelihood approach (e.g. Bauer et al. (2007) and Barwell and Schweitzer (2007)) assume that some agents are subject to DNWR, some to DRWR,

[1] In the context of more contemporary models, where productivity shocks shift the labour demand curve, the real wage rate may be procyclical.

[2] As an extreme example, in the case of firm and uniform inflation expectations, where absolutely all agents are subject to DRWR (interpreted to mean that no one will accept a real wage cut), the issue of DNWR becomes moot - except when deflation is expected. Only then will the DNWR mechanism be relevant at values of wage adjustment that exceed the expectation of inflation. Under less stringent conditions, e.g. when the anticipated inflation distribution is not degenerate and contains the point zero, it may be necessary to specify whether DNWR or DRWR takes precedence. Suppose, for instance that an agent expects inflation to be -1%, is offered a -3% wage adjustment (i.e. is subject to both a nominal and a real cut); in such a case, will the line of resistance be drawn at zero nominal adjustment (DNWR and an implied anticipated real wage increase of 1%), or at -1% nominal adjustment (DRWR and an implied real wage constancy)? Depending on how the question of which mechanism takes precedence is resolved, this will be reflected in the actual wage adjustment outcomes and the ability to distinguish the processes involved.

and some to neither type of downward wage rigidity (DWR), specifying the effects of each separately and leaving the overall picture to be determined by an endogenous mixing process. In Christofides and Nearchou (2007), we describe how the 'histogram-location' approach, that goes back to Kahn (1997), can be modified to detect not only DNWR but also DRWR. Our approach makes no parametric assumptions and does not allocate individuals to rigidity regimes, as is inherent in the likelihood-based literature. It does rely, for the identification of possible DRWR effects, on having an inflation experience which is sufficiently diverse to allow the median of the WGD to differ from the centre of the anticipated inflation distribution (AID). As the two points of central tendency drift apart, distortions to the WGD around the mean of the AID (the expected inflation rate or $\dot{P}^e$) may be detected and, if consistent with a priory restrictions, these distortions can be attributed to DRWR. DNWR is still investigated as in Kahn (1997) and Christofides and Leung (2003) by focusing on distortions in the WGD at the point zero. Thus both types of DWR can be examined. Christofides and Nearchou (2007) implement this model using wage contract data from Canada over a period (1976-1999) which is characterised by very high inflation, moderate inflation as well as extremely low inflation; however, it is applicable to any data set with a panel dimension and to any inflation period, provided sufficient care is taken in specifying the model.

In this paper, we extend this earlier work in several directions. First, the histograms are defined such that it is the median of the WGD that is located in the middle of the bin that contains it, rather than the point zero in its respective bin, as in our earlier paper. In that paper, the focus on zero allowed a neater exploration of the possibility of menu costs. Since this particular type of rigidity mechanism, while statistically significant, appears to account for approximately one percentage point of distortion in the WGD, we now wish to explore a possible lack of clarity that may arise when the median of the WGD is not centred in its so-called 'median' bin for each and every year in the sample. A second issue that we now address is the extent to which the relative frequency approach, used in our earlier work to construct the 'stage 1' bin heights that underlie our 'stage 2' econometric exploration for DWR effects, can be improved by using non-parametric kernel methods. This should be the case on *a priory* grounds, given that the relative frequency approach essentially imposes a zero bandwidth, rather than choosing one optimally. Finally, having dealt with these 'stage 1' issues, we explore the existence of and possible interactions between DNWR and DRWR by paying special attention to inflation sub-periods where (i) one kind of rigidity may be present while the other may not, as in periods of high inflation where DNWR may be not be relevant, (ii) both types of rigidity may be important but DRWR may be more important than DNWR, as in periods of moderate inflation, and (iii) both types or rigidity may be important but DNWR may be more important than DRWR, as during the more recent and prolonged period of extremely low inflation. These explorations raise technical concerns about how the model might be implemented during various sub-periods and they shed light on how these rigidities operate. They also suggest how our approach might be tailored to samples from countries and periods that share generic features with the sub-samples examined here.

Section 3.2 considers how the notions of DWR are implemented in our work; it also considers each of the three points raised in the previous paragraph, thus better-motivating the contribution of the present paper. Section 3.3 examines relevant features of the contract data that are used both here and in earlier work; working with the same data base allows useful comparisons. Section 3.4 presents the econometric specification used and its application to inflation sub-periods. Section 3.5 presents the results obtained and Section 3.6 summarises our findings and explores possible further work that might be undertaken.

3.2 Motivating Issues

In all our work this far, we assume that, *if* DNWR holds, agents would be reluctant to accept a nominal wage cut and instead would settle for a nominal wage freeze. At the population level, this reluctance would mean fewer cuts in nominal wages and more nominal wage freezes *relative to* the case of no rigidity. In terms of the distribution of nominal wage growth rates, this translates into a shift of probability mass from negative values of the support of the WGD to the point zero. Therefore the rigidity-contaminated nominal WGD would show a deficit of probability mass for negative values of the support, and a surplus at the point zero, relative to the notional distribution. At the same time, the two distributions should be identical beyond the point zero. Justifications for nominal rigidity range from the comparability and fairness arguments documented in Bewley (1999) to the theoretical papers by Macleod and Malcomson (1993), Malcomson (1997), Holden (1994), and Holden (2004) which build on the notion that nominal wages can be changed only by mutual consent.

To the extent that agents perceive that small price changes (positive or negative) are not worth the cost of implementing them, some deficit, which may not be symmetric, may appear in the area of the actual WGD immediately below and above zero. In this paper, we still check for these effects when the whole period is considered.

DRWR can be defined in a similar way to DNWR. We assume that DRWR describes the situation where agents are reluctant to accept real-wage cuts but instead would settle for a real-wage freeze. In practice, this attitude takes the form of reluctance towards accepting reductions in the *anticipated* real wage since, at the time of bargaining, future inflation is unknown. As in the case of DNWR, the presence of DRWR would distort the shape of the nominal wage growth distribution. At the population level, this would mean that agents who face nominal wage growth at a rate below anticipated inflation would settle for a nominal wage increase equal to the anticipated rate of inflation. Consequently, the presence of DRWR would shift probability mass to the right, from smaller values of nominal wage growth towards the values of anticipated inflation in the population.

The exact form of the shift of mass to the right towards the values of anticipated inflation depends on the nature of the rigidity mechanism and the joint distribution of the notional (nominal) wage growth and anticipated inflation among all agents. Nevertheless, without any distributional assumptions, it is possible to distinguish three regions in the nominal wage growth distribution for which we can make qualitative predictions about the nature of the distortions. For simplicity, suppose that the support for the AID lies inside that for the WGD. First, the interval of values that lies to the left of the support of the distribution of anticipated inflation could only loose mass to the right since all agents whose nominal wage growth falls in this region face the prospect of a real wage cut. Therefore, in this region, the rigidity-contaminated distribution can only exhibit a deficit. Second, the interval of values that lies to the right of the support of the distribution of anticipated inflation would not be distorted, since all agents whose nominal wage growth falls in this region face the prospect of a real wage increase. Third, the interval of values that corresponds to the support of the distribution of anticipated inflation, will attract mass from its left, and therefore for this interval the rigidity-contaminated distribution will exhibit a surplus in total. However, it is possible that, in some parts of this interval, the rigidity contaminated distribution will exhibit a deficit. In terms of the probability histogram, this is because a particular bin that coincides with values of anticipated inflation can attract mass from bins to its left but at the same time loose mass to bins to its right that also coincide with values of anticipated inflation. The net effect cannot be clear without knowledge of how notional wage growth and anticipated inflation are jointly distributed. The only exception is the rightmost bin in this region, for which we know that it cannot exhibit a deficit since all other bins that contain values of anticipated inflation lie to its left. Despite this uncertainty, we could assume that it would be more likely that bins that lie further to the left in this interval will show a deficit and bins further to the right will show a surplus. The sum of the net effects to the maximum point of the AID support should be zero.

This discussion indicates that the search for DRWR effects is inherently much more difficult than that for DNWR. The distortions arising from DRWR are potentially spread over a wide range of the WGD, beginning with the minimum point of the support and up to the maximum point of the AID. A further complication is that the precise limits of the support of the AID can only be conjectured; it is possible that it extends well to the left and right of $\dot{P}^e$ so that the transfer of mass may involve several bins on either side of $\dot{P}^e$. It is more likely, however, that more bins to the left will be involved than bins to the right given our discussion above.

It is also interesting to note what the presence of DRWR means for the distribution of *actual* real-wage growth. If we accept that typically the AID extends below and above the realised inflation value, then the presence of DRWR is consistent with observing real-wage cuts (relative to the realised value of inflation), even in the case of absolute (i.e. complete reluctance by all to accept a real wage cut) DRWR. Therefore, the occurrence of real-wage cuts does not, in general, suggest that DRWR does not exist; real-wage cuts are inconsistent only with the case of absolute DRWR and perfect foresight.

Having outlined our broad approach to how we expect DWR to impact on the WGD, we now turn to the three main issues that we wish to explore in this paper.

3.2.1 Centering on the Median

In our earlier work, the wage change information was used, in stage 1, to construct histograms with bins located such that the point zero was at the centre of the bin that contained it, so as to facilitate the exploration of possible menu-cost behaviour. Suppose that, in this zero-based construction, the median of the WGD was only just large enough to enter the so called 'median' bin. Then the bin containing the point zero (at its centre) might be located at the $-j$'th bin, i.e. j bins below the 'median' bin. If, on the other hand, the histograms for each year are constructed with the 'median' bin centered on the actual median for the year, then in the above example the bin containing the point zero (not at its centre) will still be the $-j$'th one. However, any other arbitrary point in the WGD support *could* belong to one bin under the zero-based construction and to an adjacent bin under median centering. An important such point is $\widehat{\dot{P}^e}$, since it figures prominently in our search for DRWR distortions. While, in practice, we do not expect these difficulties to be severe, it is preferable to now construct the yearly histograms by centering on the yearly median. The histograms presented in Figures 3.1-3.3 below are indeed constructed in this manner and are very similar with the zero-based ones presented, for selected years, in Christofides and Nearchou (2007). Note that the bins containing the expected inflation rate and the point zero are indicated in the three figures.

3.2.2 Kernel Estimates of Histogram Heights

Our test procedures involve comparisons between the notional (DWR-free) and the actual (rigidity-contaminated) WGD. These comparisons are carried out using probability histograms. We divide the support of the actual WGD into sub-intervals (bins) and compare the amount of probability mass that falls into those intervals (height of bins) to the amount in the corresponding bin of the notional. Bin width selection is driven by the nature of the data and the complexity of distortions that might be involved over intervals.

The bin heights can be formally defined as

$$P_{jt} \equiv Pr\left(\eta_{j,t} \leq \dot{w}_{ti} < \eta_{j+1,t}\right) = Pr\left(\dot{w}_{ti} \in \mathcal{B}_{jt}\right) \tag{3.1}$$

where $\dot{w}_{ti}$ is the i'th observation in year t, and $\mathcal{B}_{jt} \equiv [\eta_{j,t}, \eta_{j+1,t})$ is the j'th bin of the probability histogram in year t.

In our earlier work, we used relative frequency as the estimator of the height of bins

$$\hat{P}_{jt} \equiv \sum_{i=1}^{n} \frac{I\left(\eta_{j,t} \leq \dot{w}_{ti} < \eta_{j+1,t}\right)}{n} = \sum_{i=1}^{n} \frac{I\left(\dot{w}_{ti} \in \mathcal{B}_{jt}\right)}{n} \tag{3.2}$$

where $I(\cdot)$ an indicator function. This estimator, which could be motivated from the relative frequency definition of probability, is unbiased as well as consistent. However, in the non-parametric density estimation literature, this estimator is believed to suffer from certain problems. In particular, that it gives non-smooth estimates, that, in addition, depend critically on how the bins are defined, both with respect to their width and location.

This is the consequence of the estimator under-smoothing the data.[3]

As a robustness check, we also consider an alternative approach to estimating probability histograms that, in theory, overcomes these problems. It is based on kernel CDF estimation. To motivate the new estimator, we re-write (3.1) as follows:

$$P_{jt} = F_t\left(\eta_{j+1,t}\right) - F_t\left(\eta_{j,t}\right) \tag{3.3}$$

where $F_t(\cdot)$ is the CDF for the data in year t. Then, we get an estimator for the bin heights by plugging-in some estimator of the CDF in (3.3). The CDF estimator we consider is based on the kernel estimator of the corresponding PDF.[4] Substituting $f_t(\cdot)$ in the expression that links the CDF with the PDF by its Kernel estimator, we get

$$\begin{aligned}\hat{F}_t(\dot{w}) &= \int_{-\infty}^{\dot{w}} \hat{f}_t(u)\,du \\ &= \int_{-\infty}^{\dot{w}} \left[\frac{1}{hn}\sum_{i=1}^{n} K\left(\frac{u-\dot{w}_{ti}}{h}\right)\right] du = \frac{1}{n}\sum_{i=1}^{n} G\left(\frac{\dot{w}-\dot{w}_{ti}}{h}\right)\end{aligned} \tag{3.4}$$

where the function $G(\cdot)$ is the integral of the kernel function $K(\cdot)$. The resulting bin height estimator is then given by

$$\hat{P}_{jt} = \frac{1}{n}\sum_{i=1}^{n}\left[G\left(\frac{\eta_{j+1,t}-\dot{w}_{ti}}{h}\right) - G\left(\frac{\eta_{j,t}-\dot{w}_{ti}}{h}\right)\right] \tag{3.5}$$

This is consistent, but only asymptotically unbiased. Furthermore, it coincides with the relative frequency estimator when the bandwidth is set equal to zero.

To apply this estimator, we need to choose the type of kernel function and the bandwidth h. For the work described here we have used the Epanechnikov kernel and the Least Squares Cross Validation method to choose the optimal bandwidth.[5] This approach provides alternative estimates of the stage 1 histogram heights and we check the robustness of our stage 2 results using this alternative method of construction.

3.2.3 Tailoring the Model to Inflation Sub-Periods

The data we use are drawn from three fairly distinct periods: High inflation (1977-1982), moderate inflation (1983-1991), and low inflation (1992-1997). These should be characterised by different types of rigidity and present modeling challenges that are explored in detail in Section 3.5 below.

[3]See, for example, Silverman (1986), and Wasserman (2006) for discussion.

[4]See Li and Racine (2007).

[5]The latter decision is the most critical. The estimation was carried out in R, using the 'np' package. We are grateful to Qi Li for information and to Jeff Racine for code that implements these procedures.

3.3 Data Features

The data used in this study is derived from 10,945 collective bargaining agreements reached in all of the industries and regions of Canada between 1996 and 1999.[6] These are legally binding agreements, records for which are kept by Human Resources Development Canada (as it was known at the time the data was released to us), or HRDC. These agreements cover bargaining units involving 200 to nearly 80,000 employees, or approximately 11% of the working population of Canada in the mid year of 1989. They are derived from both the private and the public sector, and their duration ranges from a few months to several years. Because reporting requirements apply, this information is thought to be very accurate. The data set used in the empirical work below contains one observation from each of the 10,945 contracts which provides the rate of growth of the basic nominal wage rate. This growth rate refers to the total wage adjustment in the contract, including increases occasioned by the cost of living allowance clause ($COLA$). It should be noted, however, that, because the incidence and intensity of $COLA$ clauses is limited throughout the observation period, the results we obtain are similar to those that could be obtained based on non-contingent wage adjustment alone. The observation for each contract is the growth rate of the total nominal-wage adjustment over the whole of the life of the contract, calculated at *annual* rates and is allocated to the year that the contract became effective.

The data from HRDC is supplemented with information from Statistics Canada on the Consumer Price Index (CPI) inflation and an estimate of the mean anticipated inflation ($\widehat{\dot{P}^e}$) for each year.[7] These, along with the median value of the WGD for each year, appear in Table 3.1. From the CPI figures, the observation period can be divided into three consecutive periods of inflation: 1977-1983 is a high inflation period with average inflation of 9.58%; 1984-1992 is a medium inflation period with average inflation of 4.67%, and 1993-1997 is a low inflation period with average inflation of 1.46%. There is obviously a positive relationship between the yearly $C\dot{P}I$, $\widehat{\dot{P}^e}$ and the median of the realised WGD. There is also a positive relationship between the level of realised inflation, the spread of anticipated inflation and the spread of the WGDs. The latter is visually evident in Figures 3.1-3.3, where the annual histograms of the data are shown.[8] It should be noted that, within each inflation sub-period, the spread of the WGD is relatively constant.

Only 102 (or 0.9%) of the 10,945 contracts in the sample involve nominal-wage cuts, while a substantial number (1142 or 10.4%) show a wage freeze; jointly, these figures could indicate evidence in favour of DNWR. The wage freezes are particularly pronounced during the low inflation years; for each of the years 1993-1996 the proportion of contracts with a wage freeze was above 35%, peaking at 51.0% in 1993. On the other hand, 6045 (or

[6]Because of the small number of contracts involved, the first two and the last three years in the sample are considered together in everything that follows and we refer to these as ‘years’ 1977 and 1997, respectively.

[7]This is the one-year-ahead forecast from an AR(6) regression model with a GARCH(1,1) error process. This process also supplies the variance of the anticipated inflation rate at each point in time.

[8]These are median-centered, as discussed in subsection 3.4.1

55.2%) of the contracts exhibit negative real wage growth, while 4801 of them had at the same time positive nominal wage growth. These indications of real wage flexibility must be interpreted with care since they do not rule out DRWR, as has been pointed out. The number of contracts that had exactly zero real wage growth is just 1, and the remaining 4899 (or 44.8%) contracts showed both nominal and real wage increase. The econometric approach used to examine DWR is now described.

3.4 Empirical Specification

A detailed description of the econometric approach followed in our work appears in Christofides and Nearchou (2007). The basic idea is to test hypotheses about the shape of the actual-wage-growth distribution in terms of the heights of the bins of the corresponding probability histogram. We first proceed to express the actual WGDs for each year into histograms, which are then estimated non-parametrically. The resulting estimates are then used, in a second stage, in econometric estimation, where we estimate jointly the notional distribution and the distortions due to DWR.

3.4.1 Outline of the testing methodology

Testing for the presence of either type of DWR takes the form of testing hypotheses about the shape of the WGDs. Our approach is to describe the WGD with a probability histogram. Hence, the testing of hypotheses about the shape of WGDs takes the form of testing hypotheses about the height of the bins of the corresponding probability histogram.

The probability histogram for the WGD of year t could be defined as the collection of probabilities $\{P_{jt}\}_{j=-J}^{J}$, where j is the bin index. Given that our analysis focuses on the shape of the WGDs but not their location, j is defined to indicate the position of the bins relative to each other, rather than the real line. In particular, the bin indexed by $j = 0$ contains the median of the actual wage growth distribution, bins indexed by a negative j lie $|j|$ positions to the left of the median bin and bins indexed by a positive j lie j positions to its right. Furthermore, the bins of the histogram are defined such that that the median is located at the centre of the 'median' bin. We describe the probability histograms defined in this way as 'standardised', using median-centering.

In order to formulate the relevant tests, we parameterise P_{jt} under the hypotheses of no rigidity and DWR respectively by

$$P_{jt} = \begin{cases} p^N\left(z_{jt}^N; b_j^N\right) & \text{, if } H_0 \text{ is true} \\ p^R\left(z_{jt}^R; b_j^R\right) & \text{, if } H_1 \text{ is true} \end{cases} \tag{3.6}$$

where $p^N(\cdot)$ is the function of a vector of observables z_{jt}^N that gives the height of the j'th bin of the probability histogram of the notional distribution in year t, $p^R(\cdot)$ the function of observables z_{jt}^R that gives the height of the corresponding bin of the probability histogram of the rigidity-contaminated distribution in the same year, and b_j^N and b_j^R the corresponding vectors of parameters. Typically both z_{jt}^N and z_{jt}^R will contain dummy variables that will be functions of j and will indicate the relative position of bin j in the probability histogram; they may also contain additional variables that capture characteristics of the year t, while z_{jt}^R will additionally contain variables that indicate the position of bin j relative to the position of the bins containing the values taken by the rigidity bounds in the population. These variables will be functions of both j and the corresponding indices of the bins that contain the point zero (i.e. the rigidity bound for DNWR), and the anticipated inflation values (i.e. the rigidity bounds for DRWR).

With this formulation, we could test hypotheses about DWR by estimating the unrestricted model (with rigidity), and, subsequently, testing hypotheses, about the parameter vector b_j^R, that imply that the unrestricted model coincides with the restricted (rigidity free). We implement this approach in two stages:

In stage 1, the probability histogram describing the distribution underlying the observed wage growth data for each year in the sample is estimated non-parametrically. In stage 2, for each j, using the set of T estimates of the height of bin j from all years, i.e. $\{\hat{p}_{jt}\}_{t=1,\ldots,T}$, as the set of 'observations' on $\hat{P}_{jt}$,[9] we estimate the regression of $\hat{P}_{jt}$ on the vector of observables z_{jt}^R. When the estimator $\hat{P}_{jt}$ is unbiased, the regression function will coincide with $p^R\left(z_{jt}^R; b_j^R\right)$[10]. Therefore, the estimation of this equation would give estimates of the parameter vector b_j^R and its variance-covariance matrix, enabling us to test a number of restrictions related to DWR. In practice, the regression equations corresponding to all bin heights are estimated jointly since this is typically more efficient.[11]

3.4.2 Parameterisation of probability histograms

In this section we describe the most general specification of the model for the bin heights, which is estimated with the full sample. For the sub-periods, we trim this specification in order to accommodate for the special features of these periods.

Our chosen parameterisation for the heights of the bins of the probability histograms under the null hypothesis (i.e. for the notional[12] distribution), is the following

$$\begin{aligned} p^N\left(z_{jt}^N; b_j^N\right) &= \beta_{1|j|} + \beta_{2|j|} \times up_{jt} + \left(\beta_{3|j|} + \beta_{4|j|} \times up_{jt}\right) \times m_t \quad , \quad j \neq 0 \\ &= \beta_{10} + \beta_{30} \times m_t \quad , \quad j = 0 \end{aligned} \tag{3.7}$$

where m_t denotes the median of the actual-wage-growth data in year t, up_{jt} is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ lies to the right of the bin containing the median

[9] Now $t = 1, \ldots, T$ becomes the observation index.

[10] Both estimators satisfy this requirement asymptotically, and the relative frequency estimator also in finite samples.

[11] In such a case, the system would consist of $2J + 1$ equations. The dependent variable corresponding to the equation for a particular observation would be $\hat{P}_{jt}$, where j is the equation index, and t the within equation observation index. To estimate the system we would have in total $(2J + 1) \times T$ observations, with T observations on each equation.

[12] Given the parameterisation of the probability histograms under the alternative discussed below, we take the notional distribution to be the the nominal-wage-growth distribution free of any DWR or menu cost distortions.

($j > 0$), and the β's are coefficients to be estimated. With this parameterisation the $2J+1$ probability bins in each histogram can have different height from each other, therefore the notional distribution is not restricted to have any particular shape or to be symmetric. Furthermore, by making the bin height to be a linear function of the location of the actual-wage-growth distribution, and therefore of the location of the notional distribution itself, we allow for the shape of the notional distribution to vary with its location. For example, suppose that the notional distribution is symmetric around the bin containing m_t and, further, that its spread increases as its centre moves to higher values.[13] Then $\beta_{2|j|}$ and $\beta_{4|j|}$ will be equal to zero due to the symmetry assumption, $\beta_{1|j|}$ will be non-negative, and $\beta_{3|j|}$ will be negative for the bins in the middle of the distribution, i.e. for small $|j|$, and positive for the bins that lie to the tails of the distribution, i.e. for large $|j|$. Alternatively, if we allow $\beta_{4|j|}$ to be non-zero for some values of j, then the skewness of the notional distribution will also vary with the location.[14]

In order to test for the presence of both types of rigidity, the parameterisation of the probability histogram under the alternative hypothesis should reflect the distortions due to the presence of both. We assume that

$$p^R\left(z_{jt}^R; b_j^R\right) = p^N\left(z_{jt}^N; b_j^N\right) + D^u\left(z_{jt}^u; \mu\right) + D^n\left(z_{jt}^n; \gamma\right) + D^r\left(z_{jt}^r; \delta\right) \text{ , for } R = nr \tag{3.8}$$

where $D^n\left(z_{jt}^n; \gamma\right)$ is defined to be the difference between the height of the j'th bin of the rigidity-contaminated probability histogram and the height of the corresponding bin of the notional probability histogram in year t that is due to the presence of DNWR, and $D^r\left(z_{jt}^r; \delta\right)$ the corresponding difference that is due to the presence of DRWR. We also allow for distortions due to the presence of menu costs, captured by the term $D^u\left(z_{jt}^u; \mu\right)$.

For distortions due to DNWR we write

$$D^n\left(z_{jt}^n; \gamma\right) = (\gamma_1 + \gamma_2 \times m_t) \times d0_{jt} + (\gamma_3 + \gamma_4 \times m_t) \times dn_{jt} + \gamma_5 \times dz1_{jt} \tag{3.9}$$

where $d0_{jt}$ is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ contains the point zero, dn_{jt} a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is to the left of the bin containing the point zero, and $dz1_{jt}$ a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is the first bin to the right of the bin that contains the point zero. With the inclusion of the first term we can capture the distortion that applies to the bin that contains zero nominal wage growth, and, with the second term, the distortion that applies to each one of the bins that contain negative values of wage growth. In particular, γ_1 accounts for the distortion associated with the bin that contains zero nominal wage growth and γ_3 the distortion associated with the bins that lie to the left of this bin in the special case where the centre of the notional distribution, which

[13] This would imply a positive relationship between the spread and location of the histograms of the actual-wage-growth data irrespective of whether DWR is present or not.

[14] The assumption in the original Kahn (1997) methodology that the shape of the notional distribution is the same across years, has often been cited as one of the main drawbacks of this methodology as in most actual-wage-growth data sets there appears to exist a variation in the spread of the distribution across years characterised by different levels of inflation. This point is raised by Nickell and Quintini (2003) who go on to propose a flexible way of studying DNWR.

we proxy by m_t, is located at the point zero (i.e. $m_t = 0$). In that case, and, in the presence of DNWR, we would expect γ_1 to be positive, signifying the concentration of probability mass surplus in the zero nominal wage growth bin, and γ_3 negative, signifying the loss of probability mass from the bins that contain negative values of notional wage growth. When the centre of the notional distribution is located further to the right ($m_t > 0$), a smaller part of the left tail of the notional distribution lies below zero, i.e. the proportion of notional wage cuts falls, and, therefore the proportion of notional wage changes that become wage freezes due to DNWR is expected to fall. In that case, γ_2 must be negative, signifying the reduction in the probability mass surplus in the zero nominal wage growth bin, while γ_4 could be either positive or negative or zero, as the amount of mass deficit from each bin containing negative values could change in any direction relative to its level at $m_t = 0$. The inclusion of the last term enables us to test the hypothesis that, apart from shifting mass to the point of zero nominal wage growth, the presence of DNWR could also induce a shift of mass beyond the point zero, towards small positive values (in that case, $\gamma_5 > 0$) - see Holden (1989,1998,2004) and Cramton and Tracy (1992).

The distortion in the height of the probability bar of bin $\mathcal{B}_{jt}$ due to DRWR is assumed to be given by

$$D^r\left(z^r_{jt};\delta\right) = \delta_{1k} + \delta_{2k} \times J^P_t, \quad k = j - J^P_t,\ k_{\min} \leq k \leq k_{\max} \tag{3.10}$$

$$= \sum_{\nu=k_{\min}}^{k_{\max}} \left(\delta_{1\nu} + \delta_{2\nu} \times J^P_t\right) \times dp_{\nu,jt} \tag{3.11}$$

where J^P_t is the value of the index of the bin in year t that contains $\widehat{\dot{P}^e}$ (mean of AID), k is the distance between bin $\mathcal{B}_{jt}$ and that bin,[15] and $dp_{\nu,jt}$ are dummy variables indicating whether bin $\mathcal{B}_{jt}$ is located k positions from the bin that contains the centre of the anticipated-inflation distribution in year t,

$$dp_{\nu,jt} = \begin{cases} 1 & \text{if } \nu = k\left(= j - J^P_t\right) \\ 0 & \text{otherwise} \end{cases} \tag{3.12}$$

With this specification we allow for the size of the distortions to differ according to the location of the bin in the support of the anticipated-inflation distribution (through the indexing by k), and its location in the support of the notional-wage-growth distribution (through the dependence on J^P_t). In the presence of DRWR, the δ_{1k}'s, which account for the distortion when the centre of the anticipated-inflation distribution is located in the same bin as the median of the actual-wage-growth distribution ($J^P_t = 0$), are expected to be positive for the largest (and positive) values of k and negative for the smallest (and negative) values of k, signifying the shift of probability mass towards the right end of the support of the anticipated-inflation distribution. When J^P_t takes different values, the values

[15] The index k is assumed to take values from the set $\{k_{\min}, \ldots, 0, \ldots, k_{\max}\}$. The bin for which $k = 0$ contains the centre of the anticipated-inflation distribution, bins with positive values of k are located to the right of this bin, and bins with negative values to its left. The values taken by k_{min} and k_{max} are determined empirically.

of the δ_{2k}'s must be such that the distortions ($\delta_{1k} + \delta_{2k} \times J_t^P$) are qualitatively similar to the case where $J_t^P = 0$, however no specific statements can be made about their sign or size unless specific assumptions are made about the nature of the joint distribution of the notional-wage growth and anticipated inflation, and the rigidity mechanism.

Finally the effect of menu costs is parameterised as follows

$$D^u\left(z_{jt}^u; \mu\right) = \mu \times dnp1_{jt} \tag{3.13}$$

where $dnp1_{jt}$ is a dummy variable that is equal to 1 if bin $\mathcal{B}_{jt}$ is either one position to the left or to the right of the bin that contains the point zero. Therefore we allow for a symmetric loss of mass ($\mu < 0$) around and close to zero.

For the identification of the parameters of the model it is required that each type of rigidity distort different parts of the wage-growth distribution at least for some of the years in the sample. In this way, there will be sufficient variation in the dummy variables that indicate the bins that are affected by the distortions, so that these will not be collinear with the dummy variables that indicate the position of the bins in the notional probability histogram. This identification strategy is most relevant to the whole sample period where a rich inflation experience can be found. Where sub-periods are concerned, it is important to keep in mind the unique features of the period and to modify the identification strategy.

3.4.3 Estimation

For the estimation of the probability histograms in stage 1 we consider two alternative estimators; the relative frequency estimator, described by equation (3.2), and the kernel based estimator, described by equation (3.5).

Regarding the estimation in stage 2, the exact algebraic expression for the covariance between any pair of estimators that correspond to bins from the same or different probability histograms was derived, for the relative frequency case of stage 1, in Christofides and Nearchou (2007). This allows for the preferred estimator (FGLS) to be implemented but we also reported results based on Ordinary Least Squares (OLS) and the correct covariance matrix (corrected OLS). The relative frequency and Kernel approaches produce stage 1 data that are very similar indeed. As a result, the stage 2 parameter estimates are also very similar, and because of space limitations, they are not reported. Details are available on request.

3.5 Results

3.5.1 Whole Sample Results

As the first step, we implemented the model in Christofides and Nearchou (2007) but using the median-centered data discussed above. The results obtained are so similar that, in the interests of economy, are not presented here. In what follows, we always, therefore, use the median-centered data.

A natural next question is whether improvements to the specification of our earlier work can be achieved, given the new median-centered data. Small improvements are possible. In Table 3.2, we present results for the whole sample, median-centered data, based on FGLS and Corrected OLS. We have attempted to achieve parsimony in the specification for the effects of DRWR in the area to the right of the bin containing $\widehat{\dot{P}^e}$ because our variance estimates in column 5, Table 3.1 suggest that the AID is quite tight and because we want to maintain some degree of comparability with specifications for the subperiods. Note that, since DRWR could shift mass from below the minimum point in the support of the AID, similar parsimony to the left of the bin containing $\widehat{\dot{P}^e}$ is not desirable.

It is clear that all the qualitative features of the earlier paper are present. DNWR is clearly present. When the median of the WGD is zero, this type of rigidity accounts for an accumulation of nearly 9.88 percentage points of mass at the point zero and for a reduction of 1.58 points of mass in each of the bins involving negative wage growth (FGLS). The spike at zero becomes smaller as the median increases; if it were to increase to 4% (the approximate value of the median for the sample as a whole), the additional spike at zero would be 3.72 percentage points ($9.88 - 1.54 \times 4$). It will be seen below that these whole-sample estimates of DNWR average substantially higher effects during the low inflation period with lower effects in the other subperiods. The distortions due to DRWR are well-defined and in line with our expectations: When the 'median' bin also contains $\widehat{\dot{P}^e}$, that bin attracts 3.98 percentage points of additional mass and its adjacent bins about 1.7 percentage points of additional mass (FGLS results), as an approximately equal mass gets shifted from bins further to the left to the bins mentioned above. Again, these are average effects for the whole sample. These results are modified by the further interactions and menu cost effects that are allowed for and the rigidity contaminated and notional distributions (FGLS) appear more clearly in Figure 3.4, for selected years.

The apparent ability of the model to pick up the distortions occasioned by DWR is, to a large extent, due to the rich inflation experience present in the whole sample. It is, now, of interest to see how this model may be applied to the sub-periods. Since these involve fewer observations, it will be necessary to both simplify the model but also to adapt it to suit the needs and challenges of the sub-periods. We simplify by generally omitting consideration of menu cost behaviour ($\mu = 0$), of the effect discussed by Holden (1989,1998,2004) and Cramton and Tracy (1992) ($\gamma_5 = 0$), of changes in the notional distribution as the median changes ($\beta_{3|j|} = \beta_{4|j|} = 0,\ \forall j$), and of changes in the DRWR-induced distortions that may occur as the actual WGD and AID shift around ($\delta_{2k} = 0,\ \forall k$). The latter assumptions are justified by the fact that these shifts are necessarily more limited within the sub-periods. These changes reduce substantially the number of parameters that must be estimated during the sub-periods.

3.5.2 Sub-Samples

High Inflation (1977-1982)

During the period 1977-1982,[16] the WGD for the individual years does not often involve negative wage change and DNWR may not be relevant. Although DRWR distortions are the only ones that should be expected during high-inflation periods, their identification in practice may be difficult if these periods are short in duration (yielding a small number of yearly samples) and the difference between $\widehat{\dot{P}^e}$ and the median of the WGD is not sufficiently rich. To the extent that any distortions can be identified, these are, most likely, due to DRWR. Table 3.3 presents FGLS and Corrected OLS results for a version of the model that was simplified as described in the previous section. The parameter γ_1 for DNWR is not significantly different from zero, as one would expect. While the shift of mass towards the bin containing $\widehat{\dot{P}^e}$ is, to an extent, apparent, these distortions (δ_{11} to δ_{-18}) are not generally significant. This may be due to the limited number of observations involving diverse points on the WGD. We simplify the model further by imposing symmetry on the notional distribution, an assumption that has been used in several earlier papers. Table 3.4 shows that significant shifts in mass occur from several points to the left of the bin containing $\widehat{\dot{P}^e}$. The gains in the bin containing $\widehat{\dot{P}^e}$ are statistically significant in the case of the Corrected OLS results, with a gain in the bin containing $\widehat{\dot{P}^e}$ equal to 1.73 percentage points. Figure 3.5 plots the estimated (FGLS) probability histograms for the notional and actual WGDs based on Table 3.3, and Figure 3.6 the corresponding histograms based on Table 3.4.

Medium Inflation (1983-1991)

During this period, the WGD in our data extends into the negative orthant in every year of the sample. The mass of the actual WGD which is at, or below, zero is as low as 0.7% in 1988 and as high as 11.2% in 1991. At the same time, $\widehat{\dot{P}^e}$ ranges between 3.81% in 1985 and 6.05% in 1983, substantially above the point zero - see Section 3.3 above. Thus, a sizeable distance between the relevant ranges for DNWR and DRWR in the WGD exists and a clear separation and identification of the two processes may be possible.

Table 3.5 reports a parsimonious version of the model which retains asymmetry in the notional distribution. This model groups the effects on the second to eighth bin below the bin containing $\widehat{\dot{P}^e}$, thus simplifying the estimation (parameter δ_{-128} refers to this group). The results suggest that both kinds of rigidity are at work. For instance, when the median is 4% (its approximate value in this period), the additional mass in the bin containing the point zero is 2.84 percentage points ($11.36 - 2.13 \times 4$), while that in the bin containing the median and the point $\widehat{\dot{P}^e}$ is 6.62 percentage points.

It is also interesting to explore the importance of misspecifying the estimating equation, as, for instance, when DRWR is ignored while searching for DNWR, as was done in the early

[16]The sub-periods are defined with respect to the values of the estimated mean anticipated inflation, as it is the AID that determines the nature of distortions due to DRWR.

papers in this sub-literature. Table 3.6 and 3.7 show versions of the model which contain only DNWR and only DRWR respectively. While special cases appear to be successfully implemented, the quantitative effects are somewhat different from those in Table 3.5, where no exclusion restrictions are imposed. The spike at zero in Table 3.5 is underestimated by 2.54 percentage points while the shift of mass towards the bin containing the expected inflation rate in Table 3.6 is overestimated somewhat. These particular results suggest that omitting consideration of DRWR leads to bias and underestimation of the DNWR effects. Figure 3.7 plots the estimated (FGLS) probability histograms for the notional and actual WGDs based on Table 3.5.

Low Inflation (1992-1997)

When the median of the WGD is close to zero (as in 1993 and 1994), the extent to which separate DNWR and DRWR can be identified is unclear. The model must be calibrated to avoid undue overlap between DNWR-dedicated and DRWR-dedicated dummy variables. Allowing for too many bins may be inappropriate and it is necessary to also explore the possibility of finer binning. We have, therefore, redesigned the stage 1 data to allow for 0.5 percentage point wide bins so as to have a better chance of capturing the detail between the point zero and the rather low values of anticipated inflation (at most 2.24 in 1995, Table 3.1). We have also allowed for a more flexible specification for the distortions to the bins with negative values, by replacing the dummy variable dn (equation (3.9)) with bin specific dummies $dn\zeta$ ($\zeta = 1, \ldots, 6$), where ζ indicates the position of the bin to the left of the bin that contains point zero.[17] Estimates appear in Table 3.8. The astonishing concentration of mass at the bin containing the point zero is evident in the estimate for the additional height in that bin (this is also the median bin in 1993 and 1994). This bin attracts 36.01 points of additional mass. The DRWR mechanism can also be identified. Mass is shifted from points in the left of, to points near and at the bin containing $\widehat{\dot{P}^e}$. For instance, the extra mass at $\widehat{\dot{P}^e}$ is 1.88 percentage points (FGLS). Table 3.9 suggests that if, as would have been the case in the early years of this literature, a model were fitted for DNWR only, the results (Table 3.9) would credit to DNWR concentration of mass that in the more general specification of Table 3.8 belongs to DRWR, thereby overestimating DNWR by about two percentage points. Thus, suppressing the DRWR mechanism, given its statistical significance, leads to bias and is not advisable.

Figure 3.8 plots the estimated (FGLS) probability histograms for the notional and actual WGDs based on Table 3.8.

Further comments

The estimated notional WGDs appear to be different in periods characterized by different inflation levels. Most notably, they appear to be more spread in periods of higher inflation. This result could be justified based on the view that (see, for example, Friedman (1977)),

[17] The corresponding coefficients are γ_{31} to γ_{36}.

in periods of high inflation agents are prone to make bigger expectation errors regarding future levels of inflation, therefore the distribution of anticipation inflation rates is more spread, and so would be the notional WGD if the notional wage growth rates reflected, primarily, the anticipated changes in the price level and changes in productivity.

Also, there appears to be a drop in probability mass (i.e. fall in bin heights) in the estimated notional probability histograms for bins located between the '0 bin' and the 'mean anticipated inflation bin', notably in the low inflation period (see Fig. 3.8). Furthermore, we observe that this is also a feature of the data, as seen in the relative frequency histograms that are based on the actual data (Figs. 3.1-3.3), as well as the histograms of the estimated actual probability histograms (in Fig. 3.8). We note that this is a feature that is consistent with the presence of DRWR, which produces a shift of mass from these bins towards bins located further to the right. In order to explain the presence of this feature in the notional histograms we have to take into account that the model that is estimated places no restrictions on the shape of the notional histogram, except that: (i) this is the same across years within the same inflation period, and (ii) that the height of the bins of the estimated actual probability histograms is the sum of the estimated height of the corresponding notional bin and estimated size of the net rigidity distortion on that bin. Then, given that the feature in question is a true feature of the data that is also captured by the estimated model, there can be two explanations for this also appearing in the notional histograms: (1) that its presence reflects a genuine characteristic of the shape of the notional WGD, that moves away, though, from the preconception that this is a unimodal distribution, and, (2) in view of point (ii) in particular, that the model estimated is not well identified from the available data, and, that both the estimated notional bin heights and size of net distortions are biased. Indeed, a negative estimate for the size of net distortion for the bins in question (instead of positive, as it is now), and taller notional heights that do not exhibit the feature in question, would produce similar estimates for the actual bin heights. At the same time, such estimates would be consistent with the presence of DRWR, and, therefore, there would be no change in the qualitative result based on the current estimates, that are also consistent with the presence of DRWR.

3.6 Conclusion

In this paper, we explored several improvements to the method of constructing the stage 1 histograms that underly the estimation, in stage 2, of DNWR and DRWR distortions. The conceptual improvements (centering on the median of the WGD and using Kernel methods) produce stage 1 data that are very similar to data from the relative frequency approach. In the stage 2 sub-period estimations, the model performed as expected, failing to find DNWR in the high-inflation period but confirming the existence of distortions due to DNWR and DRWR in the medium and low inflation periods. An interesting issue is whether the DRWR mechanism, present in the medium and low inflation periods, would suggest that its omission (as in earlier studies) would qualify the results obtained for

DNWR. Suppressing DRWR does, indeed, modify the estimates for the spike at zero under DNWR. It is underestimated in the medium and overestimated in the high inflation period, confirming that omitting important variables is not advisable. Of course, other estimates of the DNWR mechanism (e.g. how the spike at zero diminishes as the median of the WGD increases) are also biased when the importance of DRWR is suppressed.

A particular challenge has been the identification of DRWR during the high inflation period. Our method, being data-driven and, essentially, semi-parametric, relies on there being sufficient differentiation in the relation between the median of the WGD and the mean of the AID. This may be one reason why our estimates are not well-identified. A related point is that the mass that, due to DRWR, is shifted towards the centre of the AID is larger if the expected inflation rate is high relative to the centre of gravity of the WGD. Table 3.1 shows that this is only true of one of the years in the 1977-1982 period, thus limiting the quantitative significance of DRWR. Finally, this being the period of adverse oil price shocks, may explain why more moderate wage growth would have been acceptable. Clearly more research in these important issues is warranted.

Year	#	$Med(\dot{w}_t)$	CPI	$\widehat{\dot{P}^e}$	$\widehat{Var}\left(\widehat{\dot{P}^e}\right)$
1977	226	8.20	7.55	7.22	0.4217
1978	673	7.43	8.01	8.42	0.4037
1979	569	10.11	8.95	8.45	0.3680
1980	520	11.95	9.13	9.28	0.3307
1981	450	13.10	10.16	11.66	0.3120
1982	562	10.69	12.43	10.43	0.2737
1983	643	5.00	10.80	6.05	0.3342
1984	676	4.00	5.86	4.50	0.3357
1985	519	4.04	4.30	3.81	0.3185
1986	551	4.10	3.96	4.08	0.2682
1987	557	3.83	4.18	4.37	0.2311
1988	556	4.89	4.34	3.97	0.1919
1989	493	5.22	4.05	4.83	0.1236
1990	547	5.77	4.99	4.55	0.1282
1991	530	4.19	4.76	5.91	0.4946
1992	632	2.00	5.62	1.49	0.7411
1993	516	0.00	1.49	2.00	0.4902
1994	471	0.00	1.86	0.50	0.4740
1995	460	0.68	0.16	2.24	0.4620
1996	448	0.87	2.16	1.43	0.4299
1997	346	1.87	1.62	1.95	0.3528
Total	10945				

Table 3.1: Descriptive statistics.

Parameter	FGLS		OLS-corrected	
	Estimate	Std. Err.	Estimate	Std. Err.
β_{10}	0.3571**	0.0091	0.3055**	0.0110
β_{11}	0.1060**	0.0062	0.1963**	0.0103
β_{12}	0.0645**	0.0055	0.0883**	0.0094
β_{13}	0.0508**	0.0053	0.0788**	0.0082
β_{14}	0.0374**	0.0051	0.0513**	0.0073
β_{15}	0.0132**	0.0049	0.0535**	0.0062
β_{16}	0.0408**	0.0082	0.0518**	0.0048
β_{17}	0.0247**	0.0065	0.0538**	0.0039
β_{18}	0.0178**	0.0049	0.0560**	0.0039
β_{21}	0.1558**	0.0098	0.0345**	0.0126
β_{22}	0.0133†	0.0080	-0.0101	0.0112
β_{23}	-0.0297**	0.0065	-0.0443**	0.0094
β_{24}	-0.0359**	0.0053	-0.0414**	0.0077
β_{25}	-0.0130*	0.0052	-0.0542**	0.0063
β_{26}	-0.0395**	0.0085	-0.0543**	0.0050
β_{27}	-0.0244**	0.0065	-0.0557**	0.0040
β_{28}	-0.0158**	0.0054	-0.0567**	0.0040
β_{30}	-0.0223**	0.0010	-0.0152**	0.0012
β_{31}	0.0061**	0.0008	-0.0039**	0.0011
β_{32}	0.0085**	0.0007	0.0037**	0.0009
β_{33}	0.0029**	0.0005	-0.0013†	0.0007
β_{34}	0.0009*	0.0004	-0.0014*	0.0006
β_{35}	0.0015**	0.0004	-0.0019**	0.0006
β_{36}	-0.0018*	0.0007	-0.0032**	0.0004
β_{37}	-0.0009†	0.0005	-0.0035**	0.0004
β_{38}	-0.0006	0.0005	-0.0038**	0.0003
β_{41}	-0.0186**	0.0014	-0.0063**	0.0016
β_{42}	-0.0060**	0.0012	-0.0013	0.0014
β_{43}	0.0004	0.0008	0.0028**	0.0009
β_{44}	0.0025**	0.0006	0.0041**	0.0008
β_{45}	0.0003	0.0005	0.0045**	0.0007
β_{46}	0.0030**	0.0008	0.0055**	0.0006
β_{47}	0.0015**	0.0005	0.0048**	0.0005
β_{48}	0.0012*	0.0006	0.0046**	0.0004
μ	-0.0134**	0.0026	-0.0221**	0.0016
γ_1	0.0988**	0.0051	0.1615**	0.0102
γ_2	-0.0154**	0.0010	-0.0293**	0.0019
γ_3	-0.0158**	0.0017	-0.0603**	0.0038
γ_4	0.0002	0.0004	0.0066**	0.0004
γ_5	0.0128**	0.0035	0.0092†	0.0050
δ_{-18}	-0.0088**	0.0030	-0.0030**	0.0010
δ_{-17}	0.0049†	0.0026	0.0021	0.0018
δ_{-16}	-0.0117**	0.0036	0.0010	0.0037
δ_{-15}	-0.0168**	0.0041	-0.0034	0.0053
δ_{-14}	-0.0268**	0.0044	-0.0186**	0.0067
δ_{-13}	-0.0333**	0.0047	-0.0167*	0.0078
δ_{-12}	-0.0115*	0.0052	-0.0112	0.0078
δ_{-11}	0.0175**	0.0057	-0.0087	0.0073
δ_{10}	0.0398**	0.0054	0.0237**	0.0065
δ_{11}	0.0179**	0.0045	0.0177**	0.0053
δ_{-28}	0.0076	0.0058	0.0051**	0.0019
δ_{-27}	0.0009	0.0015	0.0006	0.0020
δ_{-26}	-0.0023†	0.0012	0.0046*	0.0022
δ_{-25}	-0.0023*	0.0012	-0.0004	0.0019
δ_{-24}	-0.0045**	0.0012	0.0005	0.0018
δ_{-23}	-0.0075**	0.0014	-0.0015	0.0021
δ_{-22}	0.0068**	0.0017	0.0064*	0.0028
δ_{-21}	0.0157**	0.0024	0.0030	0.0036
δ_{20}	0.0037	0.0030	-0.0075*	0.0037
N	357		357	

Significance levels : † : 10% $*$: 5% $**$: 1%

NB: The parameters $\beta_{1|j|}$ and $\beta_{3|j|}$ refer to the symmetric part of the notional distribution, while the parameters $\beta_{2|j|}$ and $\beta_{4|j|}$ allow this distribution to be non-symmetric. The parameter μ refers to the menu cost behaviour. The γ's capture DNWR behaviour, with $\gamma_1 + \gamma_2 m_t$ measuring the spike at zero and $\gamma_3 + \gamma_4 m_t$ the deficit in the bins that contain negative values, where m_t is the median of the actual WGD from period t. The δ's capture DRWR: when the median WGD bin also contains the expected inflation rate ($J_t^P = 0$), then δ_{10} measures the extra mass due to DRWR in that bin, with parameters $\{\delta_{-11}, \ldots, \delta_{-18}\}$ measuring distortions to bins that lie to its left, and $\{\delta_{11}\}$ the distortion to the bin that lies to its right. Please see the text for a more complete explanation.

Table 3.2: Estimation results: FULL sample.

	FGLS		OLS-corrected	
Parameter	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
β_{10}	0.1396**	0.0068	0.1419**	0.0074
β_{11}	0.1613**	0.0072	0.1615**	0.0082
β_{12}	0.1342**	0.0066	0.1348**	0.0081
β_{13}	0.0660**	0.0049	0.0647**	0.0066
β_{14}	0.0293**	0.0036	0.0310**	0.0055
β_{15}	0.0114**	0.0027	0.0262**	0.0053
β_{16}	0.0089**	0.0025	0.0089*	0.0044
β_{17}	0.0052*	0.0021	0.0041	0.0042
β_{18}	0.0055†	0.0028	0.0009	0.0031
β_{21}	-0.0311**	0.0096	-0.0426**	0.0103
β_{22}	-0.0180*	0.0090	-0.0240*	0.0103
β_{23}	-0.0035	0.0066	-0.0071	0.0083
β_{24}	0.0145**	0.0052	0.0109	0.0069
β_{25}	0.0216**	0.0041	0.0067	0.0064
β_{26}	0.0168**	0.0037	0.0161**	0.0055
β_{27}	0.0062*	0.0028	0.0106*	0.0049
β_{28}	0.0040	0.0033	0.0082*	0.0037
γ_{1}	-0.0006	0.0033	-0.0017	0.0045
δ_{-18}	-0.0019	0.0033	0.0062	0.0052
δ_{-17}	0.0065†	0.0036	0.0025	0.0038
δ_{-16}	-0.0010	0.0023	0.0021	0.0043
δ_{-15}	-0.0004	0.0024	0.0096*	0.0046
δ_{-14}	-0.0031	0.0025	-0.0093†	0.0050
δ_{-13}	-0.0046	0.0031	-0.0158**	0.0057
δ_{-12}	-0.0047	0.0036	0.0011	0.0070
δ_{-11}	-0.0047	0.0057	-0.0049	0.0079
δ_{10}	0.0086	0.0073	0.0068	0.0084
δ_{11}	0.0011	0.0074	0.0039	0.0083
N	102		102	

Significance levels : † : 10% $*$: 5% $**$: 1%

Table 3.3: Estimation results: HIGH inflation period.

	FGLS		OLS-corrected	
Parameter	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
β_{10}	0.1435**	0.0067	0.1368**	0.0072
β_{11}	0.1486**	0.0046	0.1353**	0.0052
β_{12}	0.1302**	0.0043	0.1189**	0.0048
β_{13}	0.0689**	0.0030	0.0586**	0.0034
β_{14}	0.0396**	0.0023	0.0362**	0.0027
β_{15}	0.0222**	0.0018	0.0313**	0.0025
β_{16}	0.0184**	0.0017	0.0203**	0.0021
β_{17}	0.0108**	0.0013	0.0134**	0.0018
β_{18}	0.0100**	0.0014	0.0086**	0.0014
γ_1	-0.0010	0.0032	-0.0013	0.0045
δ_{-18}	-0.0058†	0.0031	-0.0025	0.0042
δ_{-17}	0.0036	0.0033	-0.0067**	0.0021
δ_{-16}	-0.0065**	0.0017	-0.0065*	0.0025
δ_{-15}	-0.0058**	0.0021	0.0013	0.0029
δ_{-14}	-0.0113**	0.0019	-0.0143**	0.0027
δ_{-13}	-0.0138**	0.0025	-0.0166**	0.0037
δ_{-12}	-0.0133**	0.0030	0.0084	0.0053
δ_{-11}	-0.0065	0.0051	0.0058	0.0069
δ_{10}	0.0080	0.0071	0.0173*	0.0080
δ_{11}	-0.0022	0.0073	0.0086	0.0081
N	102		102	

Significance levels : † : 10% $*$: 5% $**$: 1%

NB: Symmetry ($\beta_{21} = \ldots = \beta_{28} = 0$) is imposed.

Table 3.4: Estimation results: HIGH inflation period.

	FGLS		OLS-corrected	
Parameter	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
β_{10}	0.2347**	0.0073	0.2529**	0.0086
β_{11}	0.1452**	0.0060	0.1892**	0.0085
β_{12}	0.0992**	0.0045	0.1259**	0.0075
β_{13}	0.0376**	0.0028	0.0861**	0.0074
β_{14}	0.0170**	0.0023	0.0525**	0.0069
β_{15}	0.0090*	0.0036	0.0352**	0.0067
β_{16}	0.0332**	0.0083	0.0398**	0.0065
β_{21}	0.0606**	0.0089	0.0067	0.0102
β_{22}	-0.0259**	0.0061	-0.0516**	0.0083
β_{23}	-0.0058	0.0038	-0.0533**	0.0078
β_{24}	-0.0024	0.0028	-0.0377**	0.0072
β_{25}	-0.0039	0.0037	-0.0298**	0.0068
β_{26}	-0.0295**	0.0083	-0.0357**	0.0066
γ_1	0.1136**	0.0196	0.1381**	0.0209
γ_2	-0.0213**	0.0045	-0.0211**	0.0042
δ_{-128}	-0.0045**	0.0016	-0.0387**	0.0066
δ_{-11}	0.0309**	0.0068	-0.0042	0.0090
δ_{10}	0.0662**	0.0078	0.0277**	0.0088
δ_{11}	0.0215**	0.0054	0.0197**	0.0069
N	117		117	

Significance levels : † : 10% $*$: 5% $**$: 1%

NB: DNWR and DRWR effects are allowed for.

Table 3.5: Estimation results: MEDIUM inflation period.

	FGLS		OLS-corrected	
Parameter	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
β_{10}	0.2743**	0.0058	0.2643**	0.0061
β_{11}	0.1674**	0.0049	0.1806**	0.0054
β_{12}	0.1045**	0.0040	0.0949**	0.0041
β_{13}	0.0337**	0.0024	0.0474**	0.0030
β_{14}	0.0130**	0.0017	0.0138**	0.0021
β_{15}	0.0009	0.0021	-0.0035**	0.0011
β_{16}	0.0232**	0.0075	0.0011**	0.0004
β_{21}	0.0705**	0.0082	0.0298**	0.0088
β_{22}	-0.0206**	0.0056	-0.0131*	0.0059
β_{23}	0.0006	0.0034	-0.0124**	0.0040
β_{24}	0.0019	0.0023	0.0010	0.0027
β_{25}	0.0040	0.0024	0.0089**	0.0015
β_{26}	-0.0196**	0.0075	0.0030**	0.0010
γ_1	0.0882**	0.0171	0.1381**	0.0209
γ_2	-0.0153**	0.0039	-0.0211**	0.0042
N	117		117	

Significance levels : † : 10% $*$: 5% $**$: 1%

NB: DRWR effects are suppressed.

Table 3.6: Estimation results: MEDIUM inflation period.

	FGLS		OLS-corrected	
Parameter	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
β_{10}	0.2367**	0.0073	0.2529**	0.0086
β_{11}	0.1445**	0.0060	0.1892**	0.0085
β_{12}	0.0972**	0.0045	0.1259**	0.0075
β_{13}	0.0343**	0.0027	0.0861**	0.0074
β_{14}	0.0172**	0.0022	0.0820**	0.0076
β_{15}	0.0046†	0.0026	0.0458**	0.0067
β_{16}	0.0086*	0.0037	0.0416**	0.0065
β_{21}	0.0642**	0.0089	0.0067	0.0102
β_{22}	-0.0227**	0.0061	-0.0516**	0.0083
β_{23}	-0.0020	0.0037	-0.0533**	0.0078
β_{24}	-0.0023	0.0027	-0.0672**	0.0079
β_{25}	-0.0014	0.0028	-0.0404**	0.0068
β_{26}	-0.0050	0.0037	-0.0375**	0.0065
δ_{-128}	-0.0005	0.0014	-0.0387**	0.0066
δ_{-11}	0.0348**	0.0067	-0.0042	0.0090
δ_{10}	0.0693**	0.0078	0.0277**	0.0088
δ_{11}	0.0225**	0.0054	0.0197**	0.0069
N	117		117	

Significance levels : † : 10% $*$: 5% $**$: 1%

NB: DNWR effects are suppressed.

Table 3.7: Estimation results: MEDIUM inflation period.

	FGLS		OLS-corrected	
Parameter	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
β_{10}	0.1377**	0.0085	0.1490**	0.0094
β_{11}	0.0969**	0.0054	0.1143**	0.0090
β_{12}	0.0791**	0.0043	0.1110**	0.0074
β_{13}	0.0640**	0.0038	0.0544**	0.0046
β_{14}	0.0471**	0.0034	0.0444**	0.0043
β_{15}	0.0390**	0.0033	0.0400**	0.0027
β_{16}	0.0242**	0.0025	0.0218**	0.0022
β_{17}	0.0053**	0.0016	0.0041*	0.0017
β_{18}	0.0010	0.0014	-0.0058**	0.0015
β_{21}	-0.0191*	0.0076	-0.0219*	0.0105
β_{22}	0.0417**	0.0080	0.0171†	0.0100
β_{23}	0.0063	0.0062	0.0232**	0.0069
β_{24}	-0.0069	0.0052	-0.0008	0.0057
β_{25}	-0.0174**	0.0043	-0.0025	0.0045
β_{26}	-0.0125**	0.0032	0.0069†	0.0038
β_{27}	-0.0001	0.0020	0.0012	0.0022
β_{28}	0.0024	0.0017	0.0217**	0.0025
γ_1	0.3601**	0.0113	0.3159**	0.0165
γ_2	-0.1213**	0.0074	-0.1169**	0.0109
γ_{36}	-0.0199**	0.0022	-0.0201**	0.0021
γ_{35}	-0.0260**	0.0046	-0.0315**	0.0029
γ_{34}	-0.0455**	0.0034	-0.0475**	0.0036
γ_{33}	-0.0601**	0.0041	-0.0614**	0.0044
γ_{32}	-0.0748**	0.0043	-0.0860**	0.0059
γ_{31}	-0.0965**	0.0056	-0.1121**	0.0076
γ_4	0.0246**	0.0017	0.0294**	0.0021
δ_{-15}	0.0042	0.0029	0.0054*	0.0027
δ_{-14}	0.0102**	0.0034	0.0241**	0.0071
δ_{-13}	0.0040	0.0049	0.0067	0.0066
δ_{-12}	0.0207**	0.0054	0.0085	0.0072
δ_{-11}	0.0167*	0.0075	-0.0166†	0.0086
δ_{10}	0.0188**	0.0061	0.0027	0.0071
δ_{11}	0.0193**	0.0062	-0.0028	0.0070
N	102		102	

Significance levels : † : 10% $*$: 5% $**$: 1%

NB: DNWR and DRWR effects are allowed for.

Table 3.8: Estimation results: LOW inflation period.

Parameter	FGLS Estimate	FGLS Std. Err.	OLS-corrected Estimate	OLS-corrected Std. Err.
β_{10}	0.1524**	0.0076	0.1449**	0.0086
β_{11}	0.1075**	0.0044	0.1141**	0.0070
β_{12}	0.0870**	0.0037	0.1081**	0.0069
β_{13}	0.0744**	0.0032	0.0625**	0.0041
β_{14}	0.0545**	0.0031	0.0457**	0.0042
β_{15}	0.0473**	0.0028	0.0411**	0.0025
β_{16}	0.0276**	0.0024	0.0219**	0.0021
β_{17}	0.0061**	0.0016	0.0037*	0.0017
β_{18}	0.0013	0.0014	-0.0061**	0.0015
β_{21}	-0.0185**	0.0071	-0.0186†	0.0097
β_{22}	0.0496**	0.0075	0.0178†	0.0100
β_{23}	0.0023	0.0060	0.0128†	0.0067
β_{24}	-0.0086†	0.0049	-0.0020	0.0058
β_{25}	-0.0245**	0.0039	-0.0041	0.0044
β_{26}	-0.0158**	0.0031	0.0068†	0.0038
β_{27}	-0.0009	0.0020	0.0016	0.0022
β_{28}	0.0022	0.0017	0.0220**	0.0025
γ_1	0.3838**	0.0100	0.3269**	0.0157
γ_2	-0.1375**	0.0067	-0.1147**	0.0107
γ_{36}	-0.0229**	0.0021	-0.0200**	0.0021
γ_{35}	-0.0328**	0.0043	-0.0321**	0.0027
γ_{34}	-0.0522**	0.0030	-0.0473**	0.0034
γ_{33}	-0.0683**	0.0036	-0.0606**	0.0039
γ_{32}	-0.0832**	0.0038	-0.0839**	0.0053
γ_{31}	-0.1006**	0.0043	-0.1014**	0.0057
γ_4	0.0273**	0.0015	0.0296**	0.0019
N	102		102	

Significance levels : † : 10% $*$: 5% $**$: 1%

NB: DRWR effects are suppressed.

Table 3.9: Estimation results: LOW inflation period.

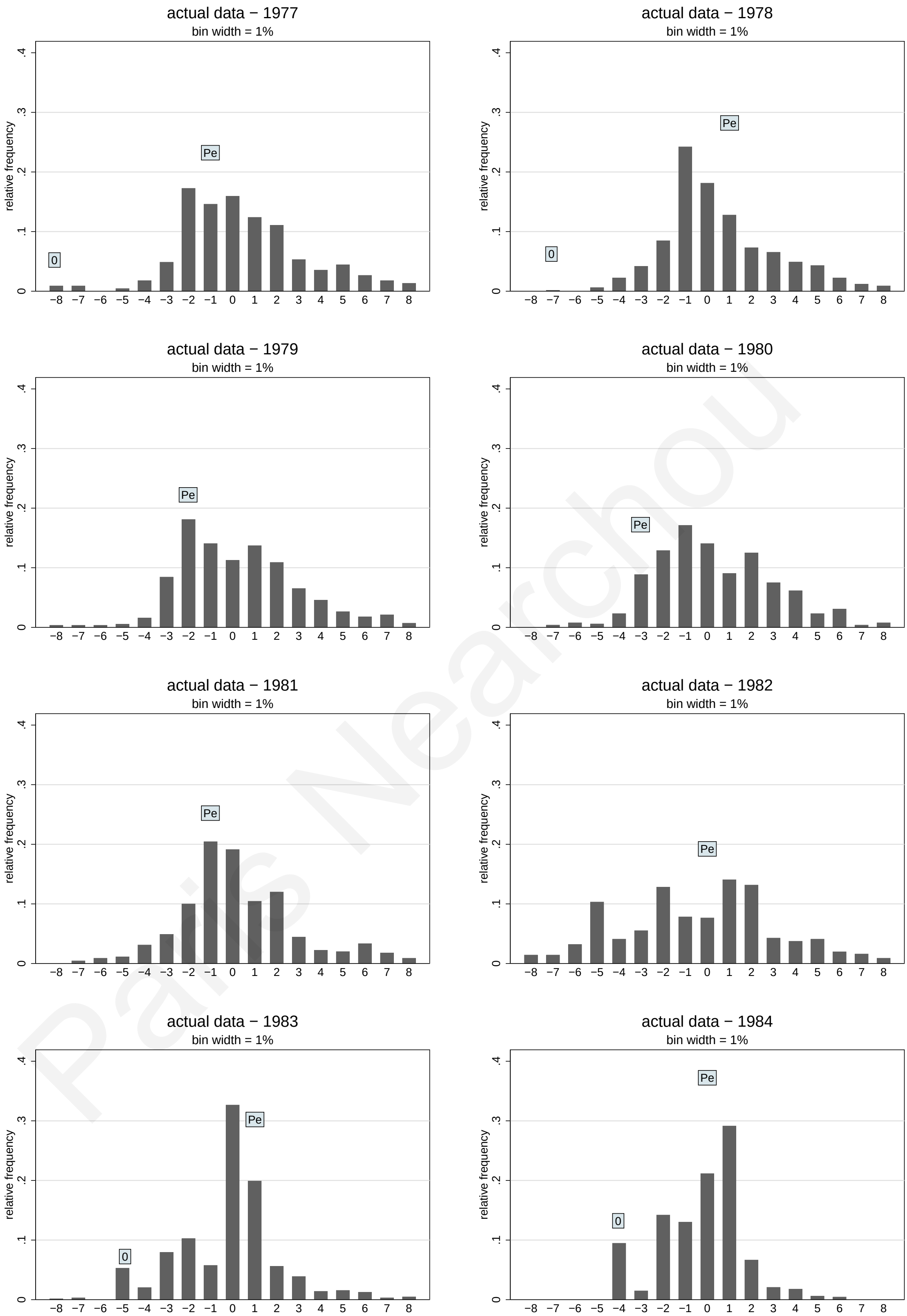


Figure 3.1: Standardised (median-centred) relative frequency histograms.

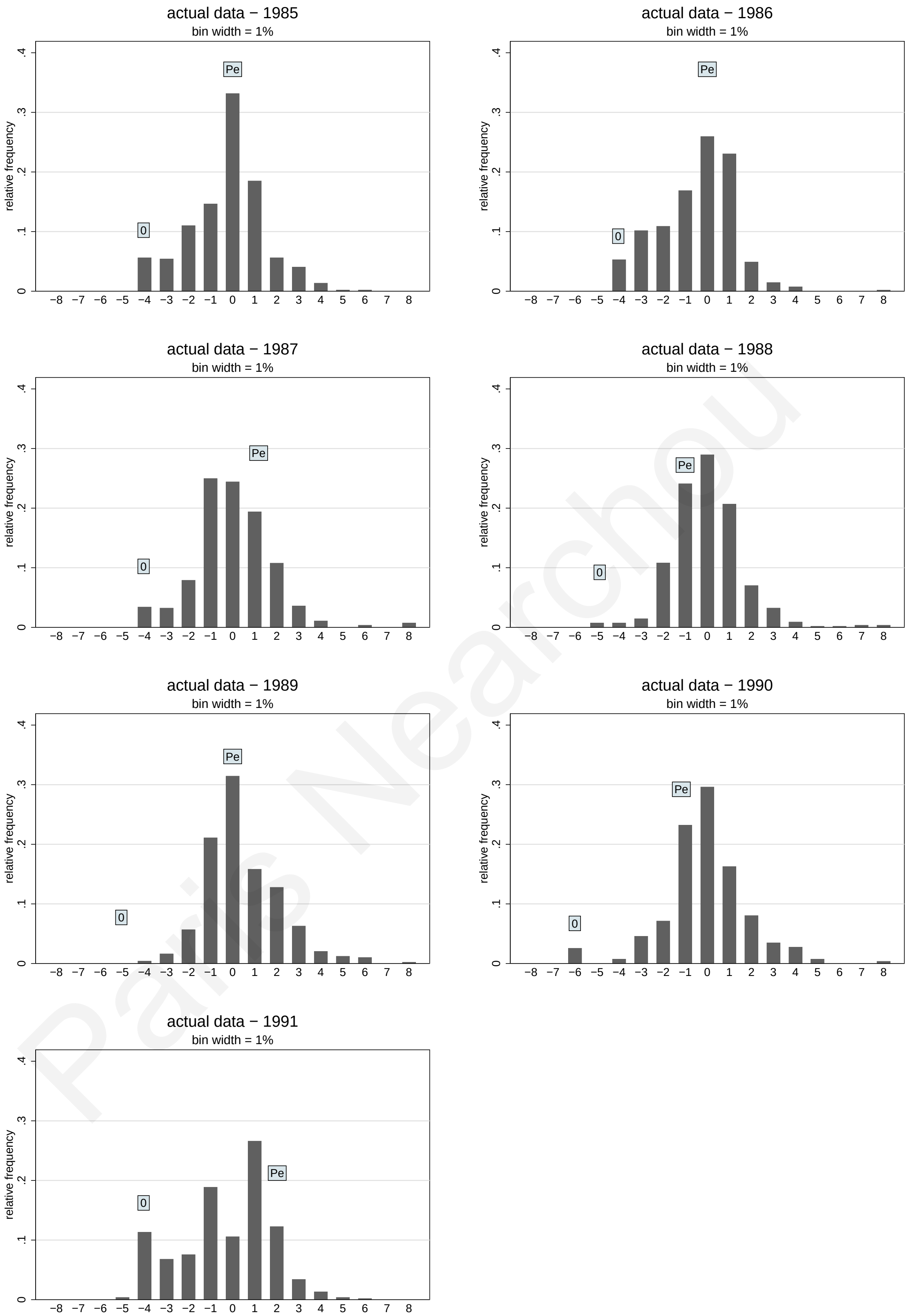


Figure 3.2: Standardised (median-centred) relative frequency histograms.

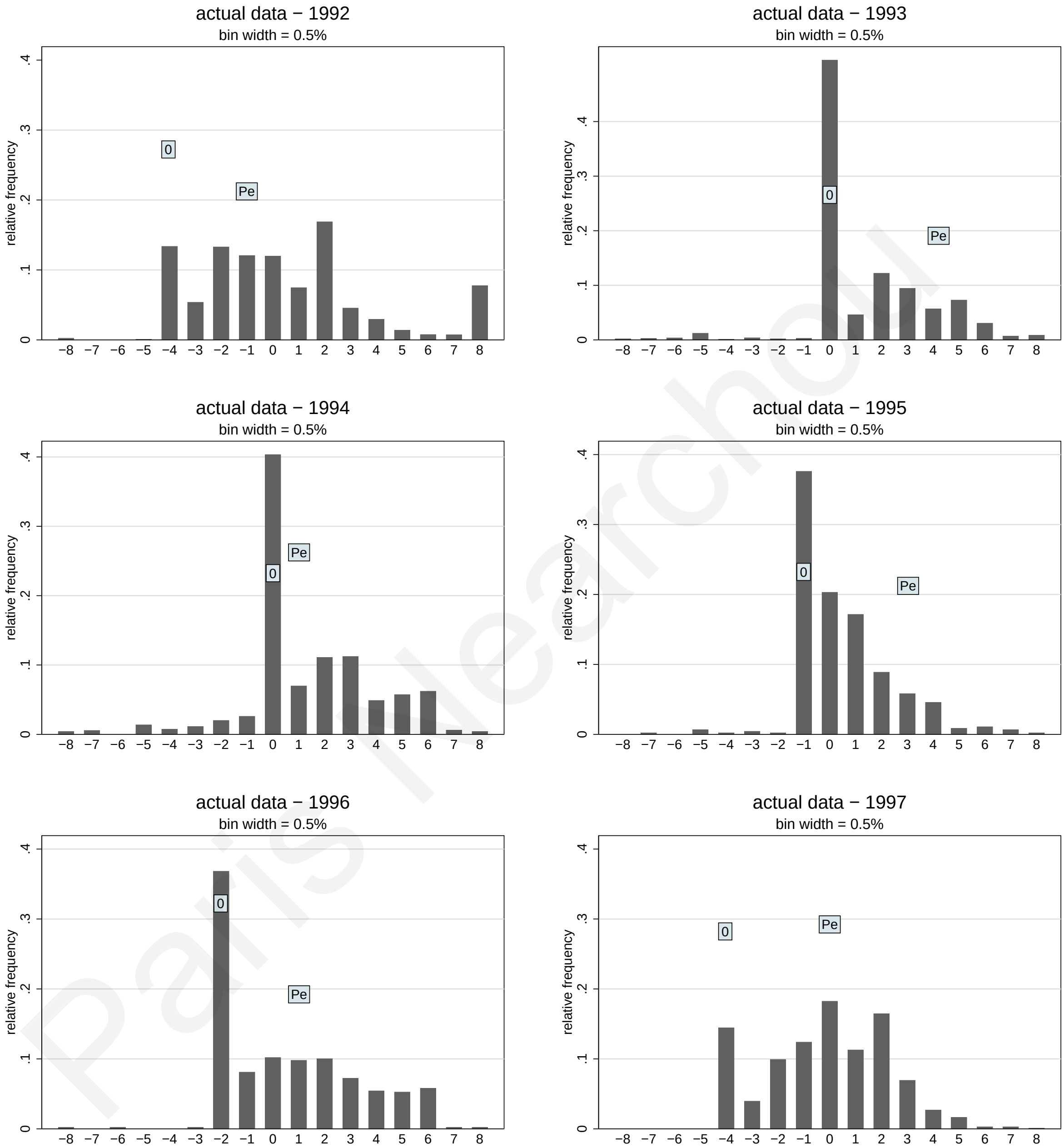
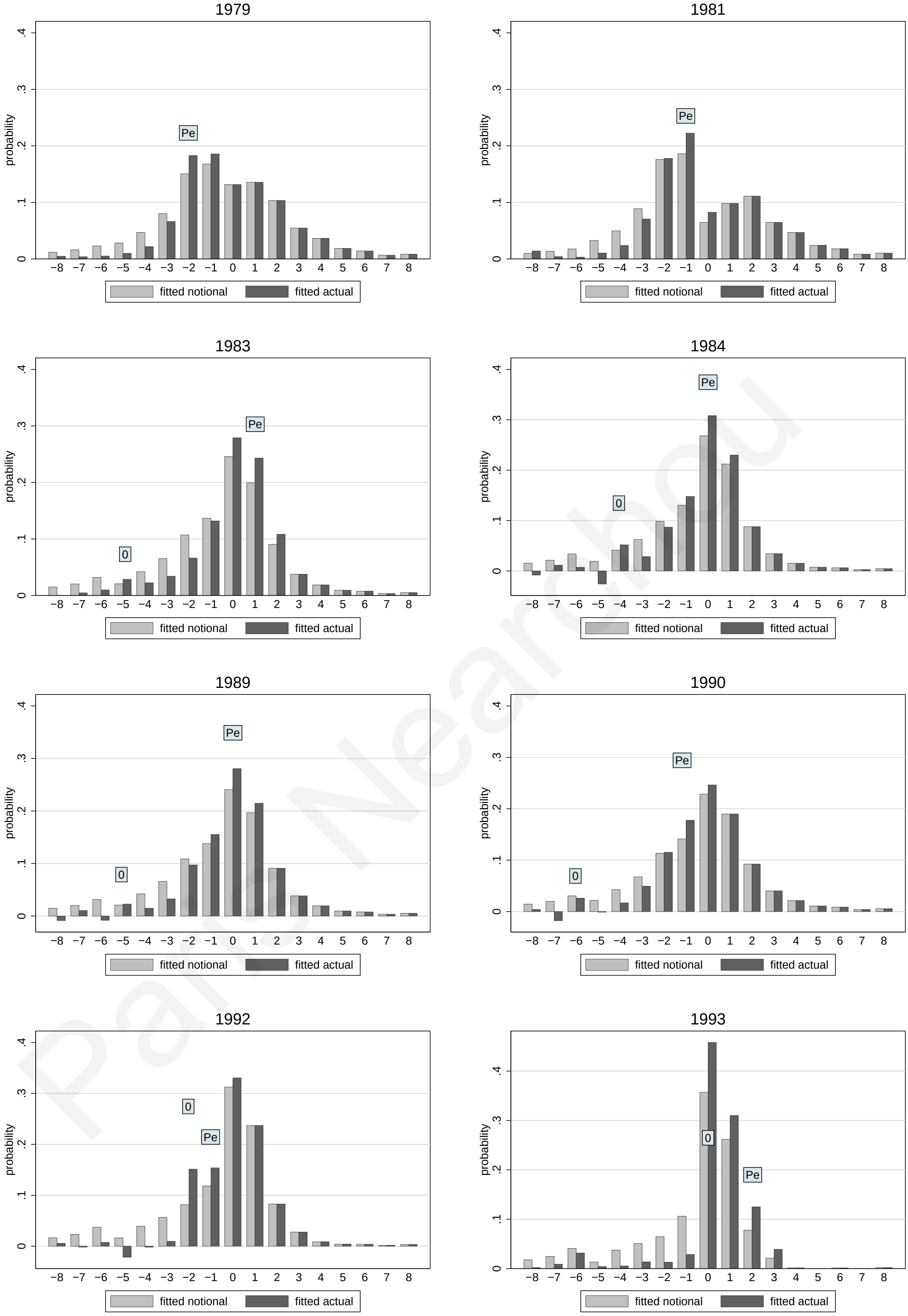
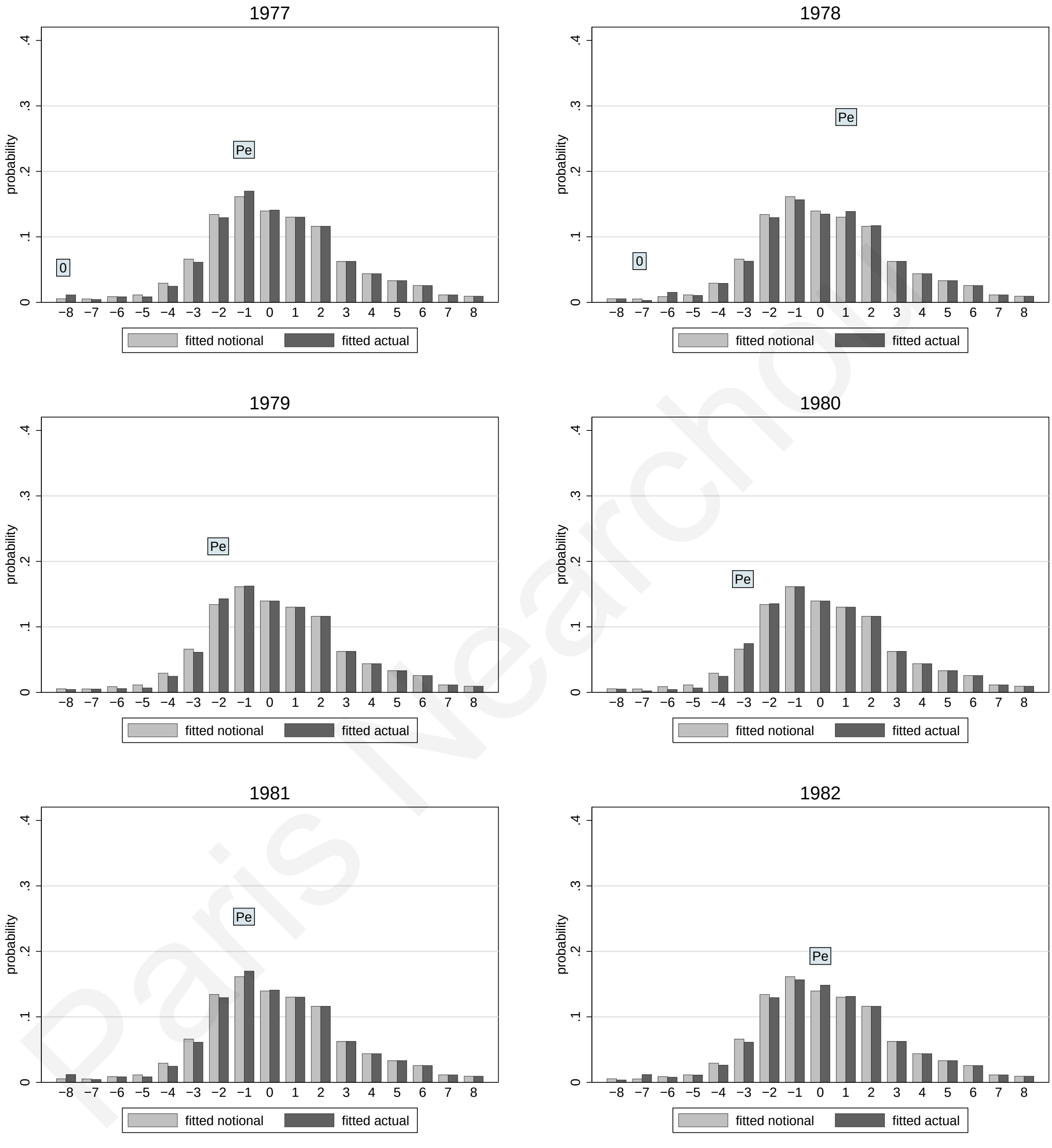


Figure 3.3: Standardised (median-centred) relative frequency histograms.



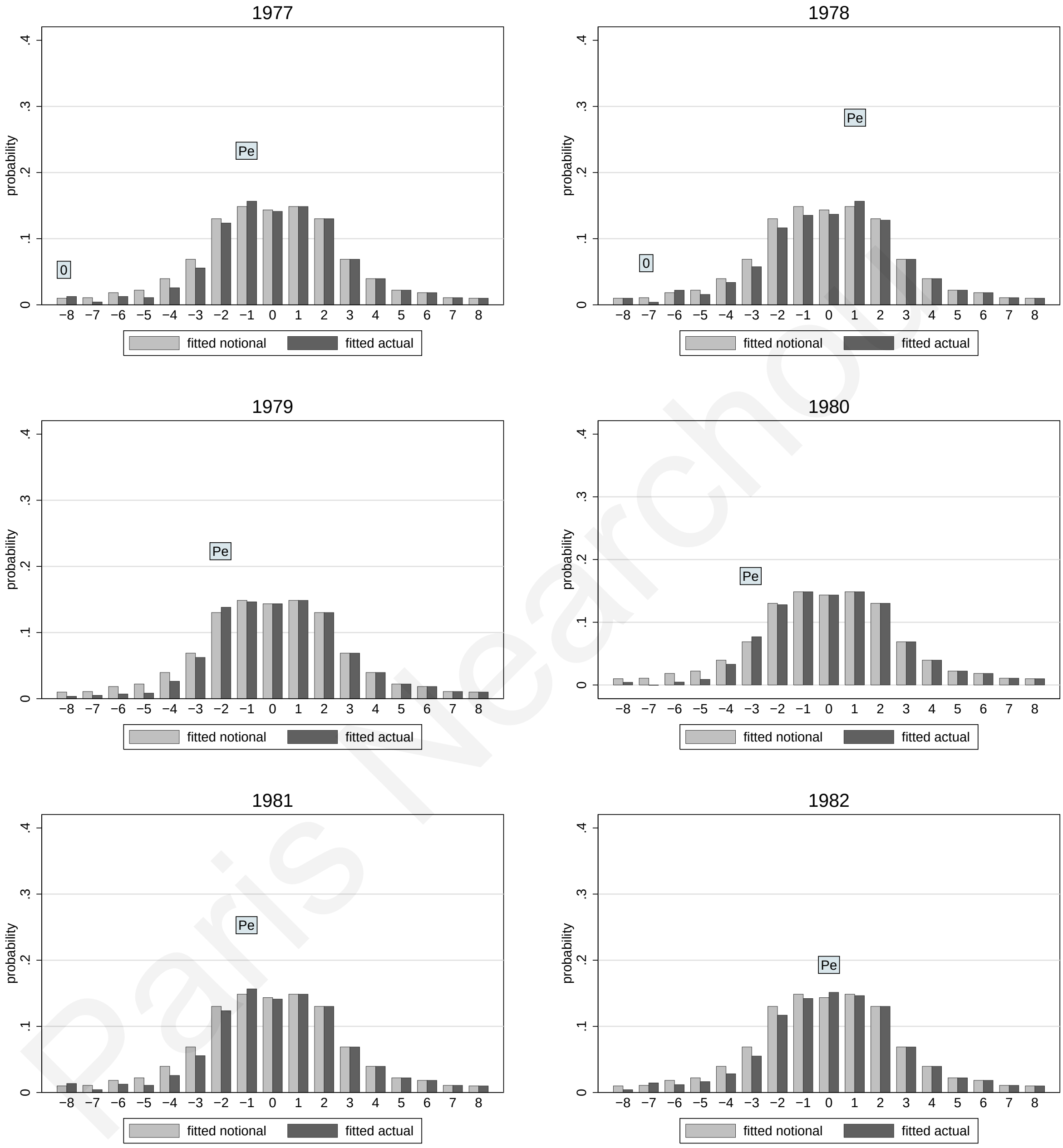
NB: FGLS results in Table 3.2 (diagrams for selected years).

Figure 3.4: Notional Vs Actual nominal WGDs (fitted values): FULL sample.



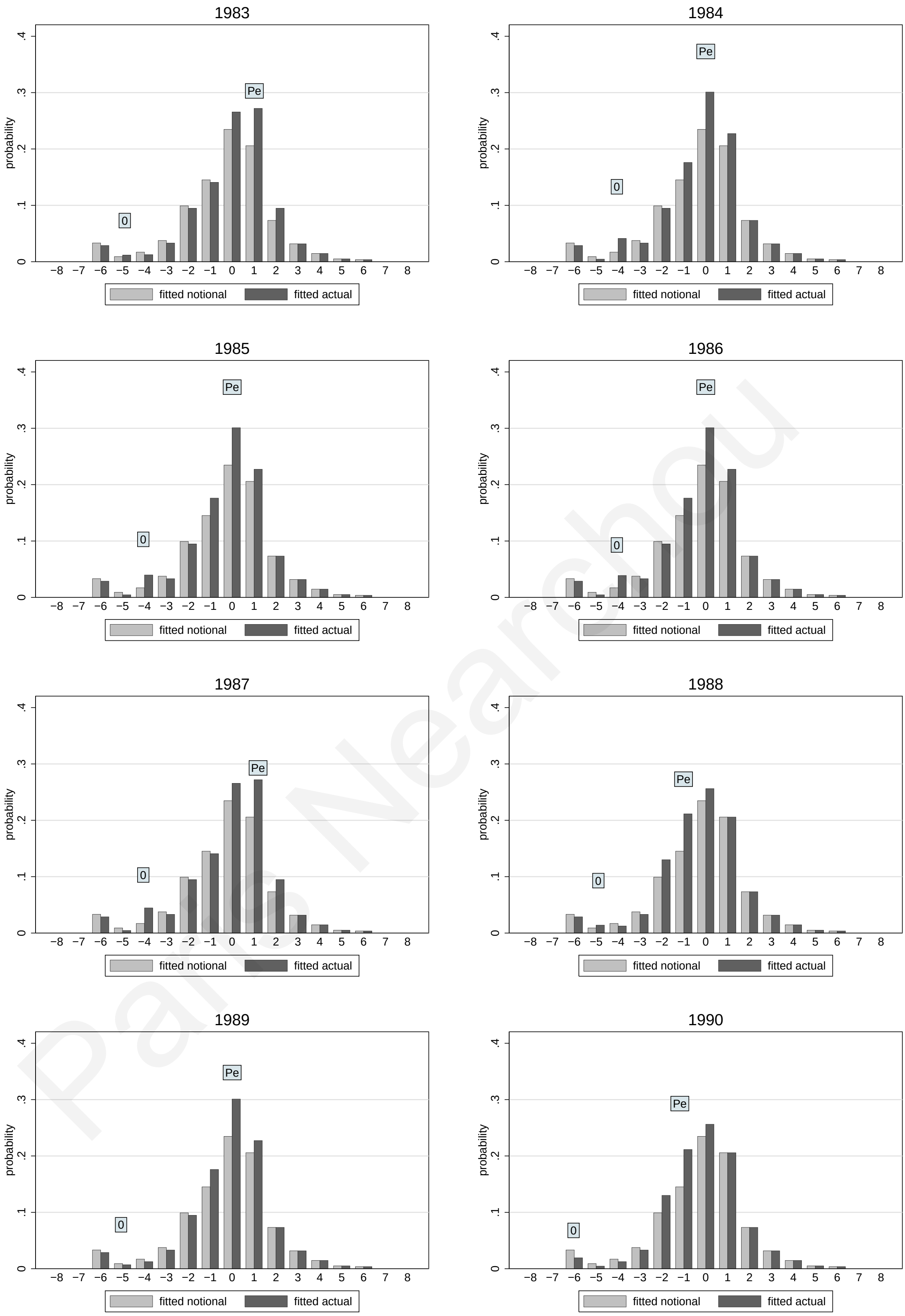
NB: FGLS results in Table 3.3.

Figure 3.5: Notional Vs Actual nominal WGDs (fitted values): HIGH inflation period.



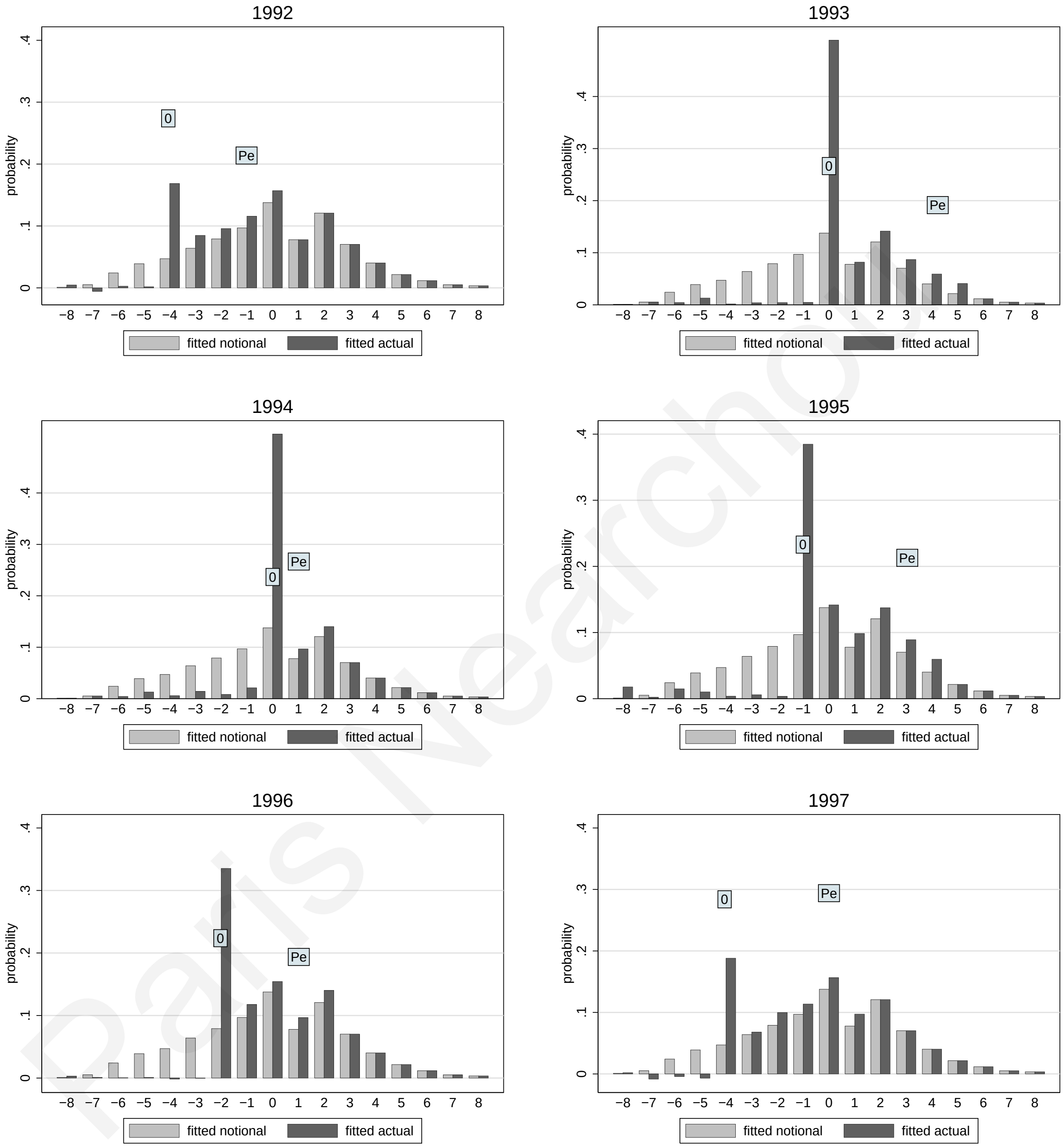
NB: FGLS results in Table 3.4 (symmetric notional).

Figure 3.6: Notional Vs Actual nominal WGDs (fitted values): HIGH inflation period.



NB: FGLS results in Table 3.5 (both types of rigidity).

Figure 3.7: Notional Vs Actual nominal WGDs (fitted values): MEDIUM inflation period.



NB: FGLS results in Table 3.8 (both types of rigidity).

Figure 3.8: Notional Vs Actual nominal WGDs (fitted values): LOW inflation period.

Chapter 4

A Semiparametric Approach for the Testing of Proportional Downward Wage Rigidity

4.1 Introduction

The issue of the existence of empirical evidence in support of the hypothesis of downward rigidity in wages has traditionally attracted considerable attention. This, of course, could be easily justified by the prominent role of the assumption of downward rigidity in nominal wages (DNWR) in macro theories since Keynes (1936), and also by the importance of the implications of these theories for economic policy.

A relatively recent example of the latter is the conjecture that DNWR induces a long-run trade-off between inflation and unemployment, and that this could provide some explanation for the persistence of high unemployment in several developed countries during the 1990's, a period that was also characterised by low inflation.[1] An implication of this conjecture is that it might be optimal for Central Banks to aim for moderate inflation levels rather than pursue a zero-inflation policy, as the benefits from moving from a near-zero to moderate inflation levels, in terms of reduced unemployment, could outweigh the costs associated with higher inflation. One can then appreciate the importance of having knowledge of the extent to which DNWR characterises the wage adjustment process in practice, and the size of the trade-off between inflation and unemployment at the estimated size of DNWR.[2]

Until relatively recently, inference about the presence of DNWR was predominantly based on evidence regarding the (counter-)cyclical behaviour of the real wage.[3] However, since the work by McLaughlin (1994), there has been a shift in the empirical work towards searching for evidence in the properties of the Distributions of Wage Growth rates (WGDs) at the micro level; in particular, looking for distortions in the *shape* of the observed WGDs that are consistent with the presence of downward rigidity.[4] A considerable number of studies have since then followed this microeconometric approach to look for evidence of downward wage rigidity using data on individual wage growth rates from a variety of sources, from several countries, and adopting a variety of strategies to identify the presence

[1] See Akerlof et al. (1996) and Holden (1994).

[2] This kind of exercise has been undertaken by several authors, including Card and Hyslop (1997) for the US, and Nickell and Quintini (2003) for the UK.

[3] Recent examples include Solon et al. (1994) and Abraham and Haltiwanger (1995).

[4] All these involve, somehow, the comparison of the shape of the observed WGD (factual) with the shape of the WGD that would prevail in the absence of rigidity (counterfactual). The latter is typically unobserved in practice, and is usually recovered from the observed wage growth data, after additional identifying assumptions are made.

of distortions.[5] Initially the focus was on the search for evidence for DNWR only, but since the early 2000's, the focus has shifted towards searching for evidence for both DNWR and Downward Real Wage Rigidity (DRWR).[6]

In this paper we describe a methodology of estimating the rigidity parameters of a model that allows for both DNWR and DRWR without making any parametric assumptions about the distributions of the observed wage growth rates. The basic idea is to approximate the continuous WGD with a discrete distribution, and then estimate a model for the probability mass function of the latter that allows for the type of distortions that could be expected in the presence of both types of rigidity. For the presentation of this methodology we adopt a simplified version of the model considered by Bauer et al. (2007), Devicienti, Maida and Sestito (2007) and Barwell and Schweitzer (2007), where wages are allowed to be formed under one of three possible wage-setting regimes; one that is flexible, and two others characterised by (essentially) proportional DNWR and DRWR, respectively.[7] We stress though, that other types of rigidity mechanisms could also be accommodated by the methodology discussed here.

The material that follows is organised in five sections: In Section 4.2 we describe a model for wage growth rates which allows for the presence of proportional DNWR and DRWR along the lines of Goette et al. (2007). Then in Section 4.3 we review the parametric approach of estimating the rigidity parameters in this model. In Section 4.4 we describe in detail the semiparametric approach that we propose here. In this Section we also discuss how this relates to the location-histogram approach, originally proposed by Kahn (1997) for the testing of DNWR. In Section 4.5 we describe several Monte Carlo simulation exercises that aim to investigate several aspects of the finite sample properties of the semiparametric estimator of the rigidity parameters, and compare these with the properties of the parametric estimator. Section 4.6 concludes.

[5]For a survey on the empirical approaches that look for evidence of DNWR see Stiglbauer (2002) and Kramarz (2001).

[6]The effort to find evidence for both types of downward wage rigidity has been spearheaded by the International Wage Flexibility Project (IWFP) - see Dickens, Goette, Groshen, Holden, Messina, Schweitzer, Turunen and Ward (2007) for a summary of their work. The interest in looking for evidence of presence of DRWR over and above the presence of DNWR is justified by the fact that DRWR can produce distortions that could be mistaken as evidence that suggest the presence of DNWR, while the policy implications from the presence of DRWR are different from those arise from the presence of DNWR. Indeed in Chapter 3 we saw that ignoring the presence of DRWR could introduce biases in the estimates of the size of the distortions attributed to DNWR.

[7]These three papers adopt the same modeling approach - summarised in Goette et al. (2007), - and investigate empirically the existence of both DNWR and DRWR in Germany, Italy, and the UK, respectively. In their case, the model is estimated after parametric assumptions are made about the distribution of the notional wage growth rates. The simplified version of their model considered here neither allows for heterogeneity in the mean of the rigidity-free wage growth rate across individuals within each year, nor for the presence of measurement error in the observed wage growth rates.

4.2 A wage growth model with proportional DNWR and DRWR

Wage-setting regimes Following Goette et al. (2007), we will assume that each wage agreement can be reached under one of three possible wage setting regimes: a flexible regime (f), where the wage growth rate is set without being constrained by any type of rigidity, one that is characterised by DNWR (n), and another that is characterised by DRWR (r). Consequently, the population of wage agreements will consist of three sub-populations, defined according to the type of wage setting regime under which the wage agreements are reached. The probability that a particular agreement belongs to a certain sub-population is, further, assumed to be fixed, and is denoted by p^f for the flexible sub-population, p^n for the DNWR sub-population, and p^r for the DRWR sub-population.

The quantity we are interested in modeling within this context is the growth rate[8] of the *nominal* wage of individual i over some period t, denoted by $\dot{w}_{ti}$, and thereon referred to as the 'actual' wage growth rate.[9] We denote by $\dot{w}^f_{ti}$ the variable that records the value of $\dot{w}_{ti}$ when it is agreed under the flexible regime, and, accordingly, $\dot{w}^n_{ti}$ and $\dot{w}^r_{ti}$ the variables that record the value of $\dot{w}_{ti}$ when it is agreed under the DNWR and DRWR regime, respectively.

The rigidity-free wage growth rate is usually referred to in this literature as the 'notional' wage growth rate, and is denoted by $\dot{w}^N_{ti}$.[10] Therefore, by construction, the actual wage growth rate set under the flexible regime is identical to the notional wage growth rate:

$$\dot{w}^f_{ti} = \dot{w}^N_{ti} \tag{4.1}$$

In the two cases where rigidity is present, the relationship between the actual and the corresponding notional wage growth rates depends on the nature of the respective rigidity mechanism. We assume that both downward nominal and real rigidity mechanisms are of the 'proportional' type, i.e. there is a fixed probability by which a notional wage change does not take place when it falls short of the relevant rigidity bound and, instead, the realised change is equal to the value of the rigidity bound. In the case of DNWR the relevant rigidity bound is the value zero, and the probability that the rigidity is binding is denoted by ρ^n. In the case of DRWR, the relevant rigidity bound is the rate of inflation anticipated by the employee side to prevail during the period the wage growth refers to (i.e. t), denoted by $\dot{P}^e_{ti}$, and the probability that the rigidity is binding is denoted by ρ^r. In all other cases, the notional wage change is realised.[11]

From these it follows that, in the case of DNWR, the relationship between the rigidity-

[8] Or, otherwise, the 'change rate'. This could be zero, or take positive or negative values.

[9] The 'individual' in this case could either be an individual employee, or an employer-union contracting pair.

[10] In practice, the notional wage growth rate is not directly observable.

[11] Formally, ρ^n is the value of the probability that DNWR rigidity is binding, conditional on, both, the event that the relevant wage setting regime for observation ti is DNWR and the particular value taken by the notional wage growth rate. The analogous definition applies for ρ^r.

contaminated and the notional wage growth rate is given by

$$\dot{w}_{ti}^{n}=\begin{cases}\dot{w}_{ti}^{N} & , \quad \dot{w}_{ti}^{N}\geq 0\\ \left.\begin{array}{ll} 0 & \text{, with probability } \rho^{n}\\ \dot{w}_{ti}^{N} & \text{, with probability } 1-\rho^{n}\end{array}\right\} & , \quad \dot{w}_{ti}^{N}<0\end{cases} \tag{4.2}$$

and, similarly, for the case of the DRWR-contaminated wage growth rate, by

$$\dot{w}_{ti}^{r}=\begin{cases}\dot{w}_{ti}^{N} & , \quad \dot{w}_{ti}^{N}\geq \dot{P}_{ti}^{e}\\ \left.\begin{array}{ll} \dot{P}_{ti}^{e} & \text{, with probability } \rho^{r}\\ \dot{w}_{ti}^{N} & \text{, with probability } 1-\rho^{r}\end{array}\right\} & , \quad \dot{w}_{ti}^{N}<\dot{P}_{ti}^{e}\end{cases} \tag{4.3}$$

Distributional assumptions For the notional wage growth rate we assume

$$\dot{w}_{ti}^{N}=\mu_{Nt}+\epsilon_{ti}^{N} \tag{4.4}$$

$$\epsilon_{ti}^{N}\sim IID \quad \text{across } i \text{ and } t \tag{4.5}$$

$$E\epsilon_{ti}^{N}=0 \quad , \; Var\left(\epsilon_{ti}^{N}\right)=\sigma_{N}^{2} \tag{4.6}$$

Therefore, the $\dot{w}_{ti}^{N}$'s are assumed to be independent of each other within and across t, and identically distributed within t, with mean μ_{Nt}:

$$\dot{w}_{ti}^{N}\overset{IID}{\sim} f_{t}^{N}\left(\dot{w}\right) \tag{4.7}$$

$$E\dot{w}_{ti}^{N}=\mu_{Nt} \quad , \; Var\left(\dot{w}_{ti}^{N}\right)=\sigma_{N}^{2} \tag{4.8}$$

Furthermore, the only form of heterogeneity across t is in the mean of the distribution, while the shape of the distribution is the same. This implies that there exists a set of values $\{\lambda_t\}$ across t, such that we could write

$$f_{t}^{N}\left(\dot{w}\right)=\tilde{f}^{N}\left(\dot{w}-\lambda_{t}\right) \quad \forall\, t \tag{4.9}$$

where $\tilde{f}^{N}(\cdot)$ is the PDF of the location-standardised notional WGD, which does not depend on t. In principle, λ_t could be any location parameter of the notional wage growth distribution, such as the mean, or any quantile of the same order across t.

We make similar assumptions about the anticipated inflation rate:

$$\dot{P}_{ti}^{e}=\mu_{\dot{P}t}+\epsilon_{ti}^{\dot{P}} \tag{4.10}$$

$$\epsilon_{ti}^{\dot{P}}\sim IID \quad \text{across } i \text{ and } t \tag{4.11}$$

$$E\epsilon_{ti}^{\dot{P}}=0 \quad , \quad Var\left(\epsilon_{ti}^{\dot{P}}\right)=\sigma_{\dot{P}}^{2} \tag{4.12}$$

therefore

$$\dot{P}_{ti}^{e}\overset{IID}{\sim} g_{t}\left(\dot{p}\right) \tag{4.13}$$

$$E\dot{P}_{ti}^{e}=\mu_{\dot{P}t} \quad , \quad Var\left(\dot{P}_{ti}^{e}\right)=\sigma_{\dot{P}}^{2} \tag{4.14}$$

Furthermore we assume independence between $\dot{w}_{ti}^N$ and $\dot{P}_{ti}^e$ across i and t.[12]

PDF of actual wage growth rate Given the assumed population structure, with the existence of three wage-setting regimes as discussed above, the (unconditional) PDF of the distribution of actual wage growth rates within this population, denoted by f_t, could be expressed as the mixture of the (conditional) PDF's characterising the distribution of wage growth rates within each sub-population:[13]

$$f_t(\dot{w}) = p^f f_t^N(\dot{w}) + p^n f_t^n(\dot{w}) + p^r f_t^r(\dot{w}) \tag{4.15}$$

where f_t^N has already been defined as the PDF of the distribution of the notional/rigidity-free wage growth rates, and f_t^n and f_t^r are the PDFs for the distributions of the wage growth rates determined under the DNWR-affected and DRWR-affected, respectively, wage setting regimes.

Using (4.2) we can derive the relationship between the probability function[14] of the DNWR-contaminated and notional WGDs;

$$f_t^n(\dot{w}) = \begin{cases} f_t^N(\dot{w}) & , \quad \dot{w} > 0 \\ f_t^N(\dot{w}) - \underbrace{\rho^n f_t^N(\dot{w})}_{loss} & , \quad \dot{w} < 0 \\ \underbrace{\rho^n F_t^N(0)}_{gain} & , \quad \dot{w} = 0 \end{cases} \tag{4.16}$$

Clearly they only coincide for non-negative values of the wage growth rate, while the magnitude of the DNWR-contaminated PDF is smaller by a proportion ρ^n for negative values. Point zero is a mass point for the DNWR-contaminated distribution, as it attracts the (non-zero measure) probability mass that is missing from the negative values of the support, which corresponds to the non-realised notional wage cuts.[15] This case is depicted in Figure 4.1, where the solid line represents the PDF of the rigidity-contaminated WGD, and the dashed line the PDF of the corresponding notional WGD.

Similarly, using (4.3), we can derive the relationship between the PDF's of the WGDs of the DRWR-contaminated and notional wage growth rates;

$$f_t^r(\dot{w}) = f_t^N(\dot{w}) - \underbrace{f_t^N(\dot{w})\left[1 - G_t(\dot{w})\right]\rho^r}_{loss} + \underbrace{F_t^N(\dot{w})\, g_t(\dot{w})\, \rho^r}_{gain} \tag{4.17}$$

[12]The stochastic independence assumptions made here are similar to those made in Goette et al. (2007). Regarding heterogeneity with respect to the mean of the notional wage growth rate, they allow for individual heterogeneity within t whereas we have assumed homogeneity. In addition, they make the parametric assumption that both the notional wage growth rates and the anticipated inflation rates are Normally distributed. Later, in Section 4.4.3, we address the issue of allowing for individual notional mean heterogeneity within t in the context of the semiparametric approach.

[13]We recall that $\dot{w}_{ti}^f = \dot{w}_{ti}^N \sim f_t^N(\cdot)$.

[14]We use this term generically, meaning either a PDF or a probability mass function. This because the actual WGD has a mass point at zero.

[15]From hereon we will refer to any differences in the shapes of the notional and actual WGDs due to rigidity as 'distortions'.

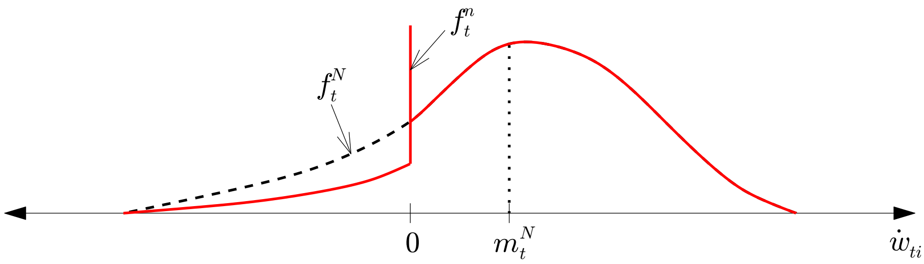


Figure 4.1: DNWR-distorted wage growth distribution.

where $F_t^N(\dot{w})$ is the cumulative distribution function (CDF) of the notional WGD and $G_t(\dot{w})$ the CDF of the distribution of anticipated inflation rates, for period t.

Expression (4.17) can become more specific if additional assumptions are made. In particular, suppose that the support of $\dot{P}^e_{ti}$ is a subset of the support of $\dot{w}^N_{ti}$,[16] i.e.

$$g_t(\dot{p}) > 0 \quad , \quad \min\left(\dot{P}^e_{ti}\right) \leq \dot{p} \leq \max\left(\dot{P}^e_{ti}\right) \tag{4.18}$$

where

$$\min\left(\dot{w}^N_{ti}\right) \leq \min\left(\dot{P}^e_{ti}\right) \quad \& \quad \max\left(\dot{P}^e_{ti}\right) \leq \max\left(\dot{w}^N_{ti}\right) \tag{4.19}$$

Then equation (4.17) can be rewritten as follows

$$f_t^r(\dot{w}) = \begin{cases} f_t^N(\dot{w}) & , \quad \dot{w} > \max\left(\dot{P}^e_{ti}\right) \\ f_t^N(\dot{w}) - \underbrace{\rho^r f_t^N(\dot{w})}_{loss} & , \quad \dot{w} < \min\left(\dot{P}^e_{ti}\right) \\ f_t^N(\dot{w}) - \underbrace{f_t^N(\dot{w})\left[1 - G_t(\dot{w})\right]\rho^r}_{loss} + & \\ + \underbrace{F_t^N(\dot{w})\, g_t(\dot{w})\, \rho^r}_{gain} & , \quad \min\left(\dot{P}^e_{ti}\right) \leq \dot{w} \leq \max\left(\dot{P}^e_{ti}\right) \end{cases} \tag{4.20}$$

This case is depicted in Figure 4.2, where the bold solid line represents the PDF of the rigidity-contaminated WGD (f_t^r) and the bold dashed curved line the PDF of the corresponding notional WGD (f_t^N). Also the thin solid line represents the PDF of the AID (g_t), while the two vertical thin dashed lines mark the bounds of the support of the AID. In this diagram we can distinguish three regions in the distribution of DRWR-contaminated wage growth rates; one to the right of the support of AID where there is no distortion (first row of the right-hand-side (RHS) of (4.20)), another region to its left where there is proportional loss of density - at a rate ρ^r - from all points (second row of RHS of (4.20)), and a third region which coincides with the support of AID (the one between the two vertical dashed lines), where, at each point, there is both loss of density towards points in

[16]This could be justified if we took the view that the change in nominal wage in the absence of rigidities, i.e. the notional wage growth rate, reflects the anticipated change in prices, as well as changes in other factors, such as individual specific productivity and demand shocks

$$\dot{w}^N_{ti} = \dot{P}^e_{ti} + \dot{\tau}_{ti}$$

where the term $\dot{\tau}_{ti}$ accumulates the effect of all other factors and has zero mean, or its mean is small relative to variance of $\dot{w}^N_{ti}$.

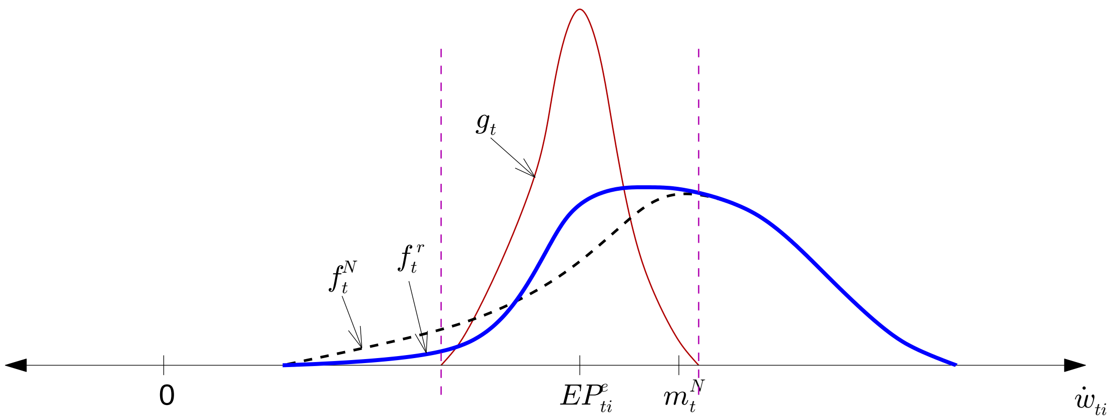


Figure 4.2: DRWR-distorted wage growth distribution.

the support of AID that located further to the right, and gain of density from points that lie to the left (third row of the RHS of (4.20)).[17] For this region, the net effect at each point will, in general, depend on the nature of the joint distribution of $\dot{w}_{ti}^N$ and $\dot{P}_{ti}^e$.

Using (4.16) and (4.17) we could, then, re-write expression (4.15) as a function of the PDF's and CDF's of the notional and anticipated inflation distributions (for the same period), and the rigidity parameters:

$$\begin{aligned} f_t(\dot{w}) = \; & \underbrace{f_t^N(\dot{w})}_{notional\ PDF} \; - \\ & - \underbrace{f_t^N(\dot{w})\, I_{(\dot{w}<0)}\; p^n \rho^n}_{DNWR\ loss} + \underbrace{F_t^N(0)\, I_{(\dot{w}=0)}\; p^n \rho^n}_{DNWR\ gain} - \\ & - \underbrace{f_t^N(\dot{w})\,[1 - G_t(\dot{w})]\;\; p^r \rho^r}_{DRWR\ loss} + \underbrace{F_t^N(\dot{w})\, g_t(\dot{w})\;\; p^r \rho^r}_{DRWR\ gain} \end{aligned} \tag{4.21}$$

where $I_{(\cdot)}$ is the indicator function that takes the value of one when the condition in the parenthesis is satisfied, and zero when it is not.

In Figure 4.3 we plot two examples, where we have assumed, for both cases, that $\min(\dot{w}_{ti}^N) < \min(\dot{P}_{ti}^e)$ and $\max(\dot{w}_{ti}^N) > \max(\dot{P}_{ti}^e)$. Therefore DRWR introduces distortions in both cases. On the other hand we have assumed $\min(\dot{w}_{ti}^N) > 0$ for the example depicted in the top diagram, and $\min(\dot{w}_{ti}^N) < 0$ for the example depicted in the bottom diagram. Therefore DNWR is only relevant in the second case, producing its trademark spike at point zero.

In the next two sections we consider the estimation of the model just described. Our primary interest is in making inferences about the rigidity parameters (ρ^n, ρ^r) and the mixing probabilities (p^n, p^r). In Section 4.3 we start off with the description of the implementation of the parametric approach, and in Section 4.4 we continue with the description of the semiparametric approach that is proposed here.

[17] We note that the case of DRWR could be seen as a generalisation of the case of DNWR if, in the latter case, we viewed the distribution of rigidity bounds as degenerate at point zero.

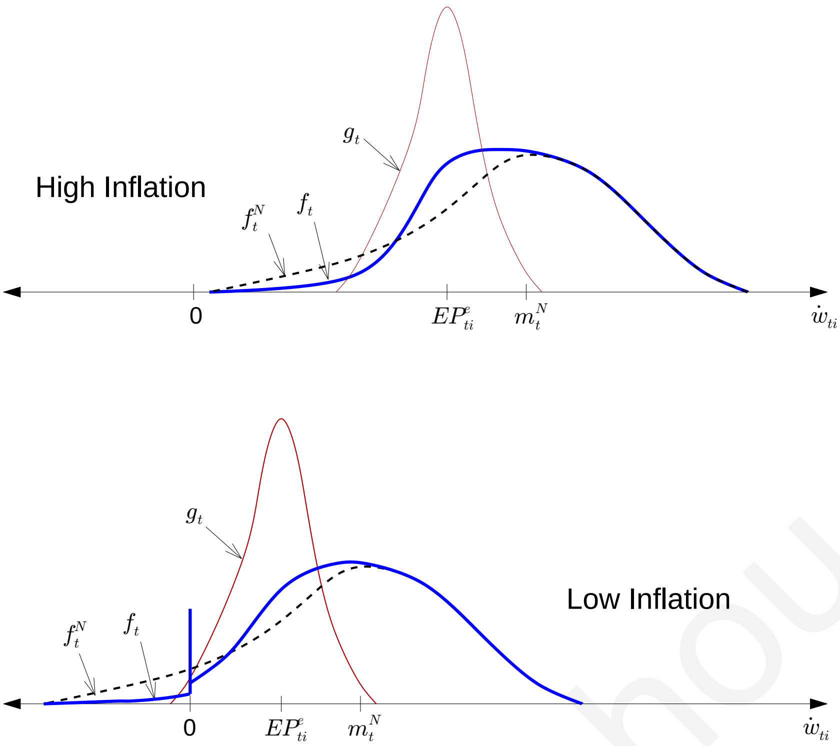


Figure 4.3: DNWR- and DRWR-distorted wage growth distributions.

4.3 Parametric estimation

Description The parametric approach to estimating the parameters of interest consists of making specific assumptions about the parametric families of the distributions of the notional wage growth rate and the anticipated inflation rate, and then estimating the parameters of these distributions along with the parameters of primary interest, using maximum likelihood.[18]

Usually we have observations on wage growth rates that extend for several years

$$\{\dot{w}_{ti}\}_{\substack{t=1,\ldots,T \\ i=1,\ldots,n_t}} \tag{4.22}$$

Typically we do not observe in which sub-population a particular wage agreement (ti) belongs to, despite that this is known, at least, to the employee side of the negotiating pair.[19] Therefore we consider the PDF in (4.21) to be the appropriate contribution to the

[18]A more general formulation of this approach appears in Goette et al. (2007). The formulation that appears here has been adjusted to accommodate the simpler features of our model.

[19]Had we observed the wage setting regime for each agreement, the appropriate contributions to the likelihood function would be given by $f_t^N(\dot{w})$ for the flexible regime, by $f_t^n(\dot{w})$ in expression (4.16) for the DNWR regime, and by $f_t^r(\dot{w})$ in expression (4.17) for the DRWR regime.

likelihood function of the observation for individual i in period t [20]

$$L_1 = \prod_{t=1}^{T} \prod_{i=1}^{n_t} f_t\left(\dot{w}_{ti}\right) \tag{4.23}$$

Empirical implementation issues To proceed with the estimation we need to make assumptions about the parametric families of the distributions of the notional wage growth rate and the anticipated inflation rate, over and above the distributional assumptions already stated in the previous section. Typically Normality is assumed:[21]

$$\dot{w}_{ti}^{N} \sim N\left(\mu_{Nt}, \sigma_N^2\right) \tag{4.24}$$

$$\dot{P}_t^{e} \sim N\left(\mu_{\dot{P}t}, \sigma_{\dot{P}}^2\right) \tag{4.25}$$

therefore the likelihood function takes the form

$$\begin{aligned} L_1 = {} & \underbrace{\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right)}_{notional\ PDF} - \\ & - \underbrace{\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right) I_{(\dot{w}<0)}\, p^n \rho^n}_{DNWR\ loss} + \underbrace{\Phi\left(-\frac{\mu_{Nt}}{\sigma_N}\right) I_{(\dot{w}=0)}\, p^n \rho^n}_{DNWR\ gain} - \\ & - \underbrace{\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right)\left[1-\Phi\left(\frac{\dot{w}-\mu_{\dot{P}t}}{\sigma_{\dot{P}}}\right)\right] p^r \rho^r}_{DRWR\ loss} + \underbrace{\Phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right)\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{\dot{P}t}}{\sigma_{\dot{P}}}\right) p^r \rho^r}_{DRWR\ gain} \end{aligned} \tag{4.26}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the CDF and PDF, respectively, of the Standard Normal distribution.

From the above expression we note that p^n and ρ^n, on the one hand, and p^r and ρ^r, on the other hand, always appear multiplied to each other in the likelihood function. Therefore the individual parameters in each pair cannot be identified separately but, instead, only their respective products, which we denote by

$$\rho^{nn} \equiv p^n \rho^n \tag{4.27}$$

$$\rho^{rr} \equiv p^r \rho^r \tag{4.28}$$

[20] As the model does not allow for the presence of measurement error in the observed wage growth rates, the observed growth rates are the same as the actual. Consequently, the above results, which refer to the *actual* growth rates, could also be used directly for the econometric analysis of the *observed* data.

[21] The assumption of Normality has been the predominant assumption for the notional distribution in the context of testing for DNWR only; by Altonji and Devereux (2000), who introduce the parametric approach, and thereafter, by Knoppik and Beissinger (2003) and Fehr and Goette (2005), among others. As we have already mentioned, Normality has also been assumed for both the notional and AID by Barwell and Schweitzer (2007), Bauer et al. (2007) and Devicienti et al. (2007) in the context of testing for both DNWR and DRWR. In all the above cases, this assumption was used in more complex setups where measurement error was also present. An alternative assumption to Normality is the generalised hyperbolic distribution, proposed by Behr and Pötter (2005) and implemented in the context of testing only for DNWR.

Each of these products could be interpreted as the 'overall' or 'total' effect of the respective type of rigidity, as they measure the size of the effect of rigidity in the entire population, whereas the individual parameters ρ^n and ρ^r measure the size of this effect within the sub-population that is affected by the particular type of rigidity.[22]

Having said that, the estimates of ρ^{nn} and ρ^{rr} can provide estimates for the lower bound of ρ^n and ρ^r, respectively. For example, in the case of ρ^{nn}, given the relationship in (4.27) and that $\max(p^n) = 1$, then it must be true that $\min(\rho^n) = \rho^{nn}$. Therefore the estimate of ρ^{nn} provides an estimate for the lower bound of the interval of values that could contain ρ^n. Similarly, the estimate of ρ^{rr} provides an estimate for the lower bound of the interval of values that could contain ρ^r.[23]

We conclude this section by noting that in this simple setup, where we have assumed mean homogeneity for the notional WGD within t, we would have to fix the mean of the AID in order to achieve the identification of the remaining parameters. Under the assumption that inflation expectations are unbiased, then the mean *anticipated* inflation for period t will coincide with the *realised* inflation rate in that period

$$\mu_{\dot{P}t} = \dot{P}_t \quad , \; t = 1, \ldots, T \tag{4.29}$$

It follows from the above that the parameters of the model left to be estimated are

$$\begin{aligned} \rho^{nn} &: \text{ Total DNWR effect} \\ \rho^{rr} &: \text{ Total DRWR effect} \\ \{\mu_{Nt}\}_{t=1,\ldots,T} &: \text{ mean of notional distributions for each period} \\ \sigma_N &: \text{standard deviation of notional distributions} \\ \sigma_{\dot{P}} &: \text{standard deviation of AID's} \end{aligned} \tag{4.30}$$

4.4 Semi-parametric estimation

In this Section we describe a method of estimating the rigidity parameters in the model presented in Section 4.2. We recall that the distributional assumptions made there did not go as far as to specify the family of distribution the actual WGD belongs to. Therefore the approach described here could be described as semiparametric.

The basic idea underlying the semiparametric approach is the approximation of the continuous distribution of the actual wage growth rate with a discrete distribution and searching for the presence of distortions, that are consistent with the presence of rigidity,

[22] If we were to adopt the relative frequency interpretation of probability, then the product $\rho^n p^n$ would give the proportion of wage agreements in the population with zero growth rate that in the absence of downward nominal rigidity would be characterised by a negative growth rate. Similarly, the product $\rho^r p^r$ would give the proportion of wage agreements in the population with growth rate equal to the relevant anticipated inflation rate that in the absence of downward real rigidity would be characterised by a growth rate lower than that.

[23] It also follows that, for the case where it is known a priori that we have absolute rigidity, i.e. $\rho^n = \rho^n = 1$, then ρ^{nn} and ρ^{rr} measure the size of coverage of each wage-setting regime in the population.

in the shape of the latter.

The implementation of the semiparametric approach involves three stages: firstly, the standardisation of the location of the actual WGDs across periods, then, the discretisation of the location-standardised actual WGDs and, finally, the estimation[24] of the model for the probability mass function (PMF) of the discrete distributions which, by construction, involves the rigidity parameters.

Next we look at the details of the implementation of each stage in turn.

4.4.1 Details of method

Standardisation The purpose of this exercise is to standardise the location of the actual WGD for each period such that the underlying notional WGDs are identical across periods.

We define the location-standardised variables

$$\tilde{w}_{ti} \equiv \dot{w}_{ti} - \lambda_t \sim \tilde{f}_t\left(\tilde{w}\right) \tag{4.31}$$

$$\tilde{w}_{ti}^N \equiv \dot{w}_{ti}^N - \lambda_t \sim \tilde{f}^N\left(\tilde{w}\right) \tag{4.32}$$

where λ_t has already been defined in (4.9) to be a location parameter of the notional WGD for period t that is chosen such that the location-standardised notional WGDs are identical across t.

Using λ_t, we also define the location-standardised versions of the variables that record the anticipated inflation rate, and the wage growth rates under the three wage-setting regimes;

$$\tilde{P}_{ti}^e \equiv \dot{P}_{ti}^e - \lambda_t \sim \tilde{g}_t\left(\dot{p}\right) \tag{4.33}$$

$$\tilde{w}_{ti}^f \equiv \tilde{w}_{ti}^N \sim \tilde{f}^N\left(\tilde{w}\right) \tag{4.34}$$

$$\tilde{w}_{ti}^n \equiv \dot{w}_{ti}^n - \lambda_t \sim \tilde{f}_t^n\left(\tilde{w}\right) \tag{4.35}$$

$$\tilde{w}_{ti}^r \equiv \dot{w}_{ti}^r - \lambda_t \sim \tilde{f}_t^r\left(\tilde{w}\right) \tag{4.36}$$

Given that the effect of the standardisation is to shift all distributions within a particular period by the same distance ($= \lambda_t$) from their original location it follows that similar relationships to those in (4.15)-(4.17) and (4.20), that hold for the original distributions, will hold for the location-standardised distributions. That is, it must be true that

$$\tilde{f}_t\left(\tilde{w}\right) = p^f \tilde{f}^N\left(\tilde{w}\right) + p^n \tilde{f}_t^n\left(\tilde{w}\right) + p^r \tilde{f}_t^r\left(\tilde{w}\right) \tag{4.37}$$

where

$$\tilde{f}_t^n\left(\tilde{w}\right) = \begin{cases} \tilde{f}^N\left(\tilde{w}\right) & , \quad \tilde{w} > -\lambda_t \\ \tilde{f}^N\left(\tilde{w}\right) - \rho^n \tilde{f}^N\left(\tilde{w}\right) & , \quad \tilde{w} < -\lambda_t \\ \rho^n \tilde{F}^N\left(-\lambda_t\right) & , \quad \tilde{w} = -\lambda_t \end{cases} \tag{4.38}$$

[24] We will consider estimation using maximum likelihood. Later in Section 4.4.5, we also discuss a regression approach.

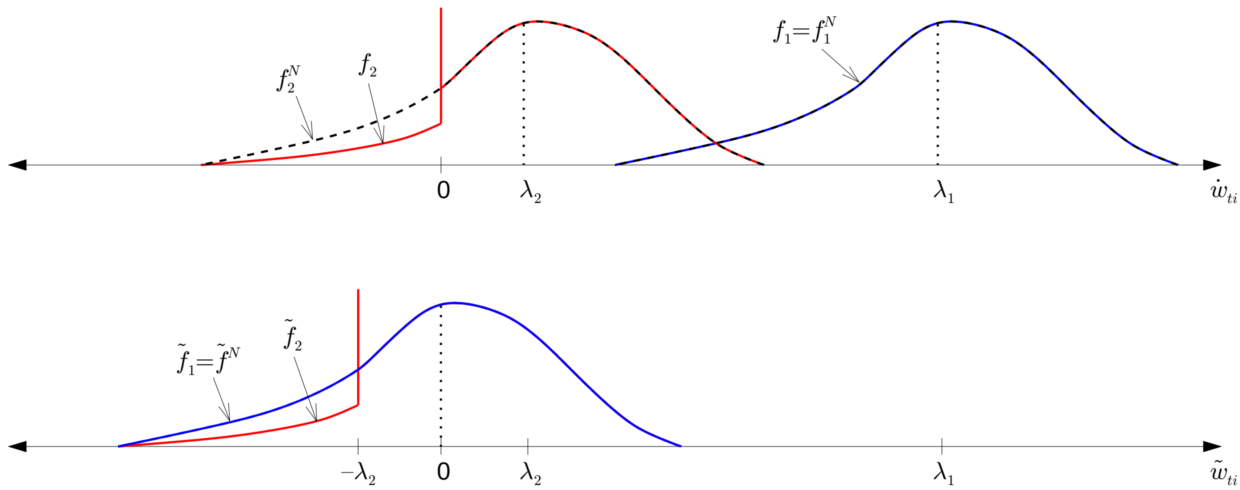


Figure 4.4: Location standardisation.

and

$$\tilde{f}_t^r(\tilde{w}) = \tilde{f}^N(\tilde{w}) - \tilde{f}^N(\tilde{w})\left[1 - \tilde{G}_t(\tilde{w})\right]\rho^r + \tilde{F}^N(\tilde{w})\,\tilde{g}_t(\tilde{w})\,\rho^r \tag{4.39}$$

$$= \begin{cases} \tilde{f}^N(\tilde{w}) & , \ \tilde{w} > \max\left(\tilde{P}_{ti}^e\right) \\ \tilde{f}^N(\tilde{w}) - \rho^r \tilde{f}^N(\tilde{w}) & , \ \tilde{w} < \min\left(\tilde{P}_{ti}^e\right) \\ \begin{array}{l} \tilde{f}^N(\tilde{w}) - \tilde{f}^N(\tilde{w})\left[1 - \tilde{G}_t(\tilde{w})\right]\rho^r + \\ +\tilde{F}^N(\tilde{w})\,\tilde{g}_t(\tilde{w})\,\rho^r \end{array} & , \ \min\left(\tilde{P}_{ti}^e\right) \le \tilde{w} \le \max\left(\tilde{P}_{ti}^e\right) \end{cases} \tag{4.40}$$

The purpose of this type of standardisation is to facilitate the comparison of the shapes of the WGDs across periods, and, ultimately, the identification of the rigidity parameters. To see how this could be achieved, we note that in the absence of rigidity the standardised actual WGDs will coincide with the corresponding standardised notional WGDs, and thus by construction (see (4.34)), be identical across t. On the other hand, in the presence of rigidity, the standardised actual WGDs from each period will be distorted, possibly in a different way, depending on the location of the notional WGD with respect to the relevant rigidity bounds; the point zero and the support of the AID. Then, the comparison of the parts of the standardised WGDs across periods that are expected to be distorted differently in the presence of rigidity would provide information about the size of the distortions that ultimately could lead to the identification of the rigidity parameters.

An example of implementation of the standardisation stage is depicted in Figure 4.4. For simplicity we present a case where only DNWR affects the wage adjustment process. The top diagram shows the PDFs of the notional (dashed line) and actual (solid line) WGDs from two periods before standardisation, and the bottom diagram the corresponding PDFs after standardisation takes place. In this example we have assumed that in period 1 the inflation level is high enough so that no part of the notional WGD lies below zero, and therefore the PDF of the notional (f_1^N) and actual (f_1) WGDs are identical for this case. On the other hand, the inflation level in period 2 is assumed to be low enough so that a part of the notional distribution lies below zero. As a result, for this case, the notional and actual

PDFs do not coincide for all levels of wage growth rates, but instead the characteristic distortion in the shape of the PDF of the actual WGD (f_2) is visible for the part that corresponds to the non-positive part of the support of the notional WGD.

Standardisation is assumed to take place using as standardisation parameters the values λ_1 and λ_2 for periods 1 and 2, respectively, which are the values of the same-order-quantiles of the unstandardised notional WGDs in the two periods. As we can see in the bottom diagram, the PDF of the standardised actual WGD in period 1 ($\tilde{f}_1$) coincides with the PDF of the standardised notional WGD ($\tilde{f}^N$), which, by assumption, is the same for both periods. On the other hand, the shape of the PDF of the standardised actual WGD for period 2 ($\tilde{f}_2$) preserves the distortion in its shape, except that instead of this extending to the left - and including - point zero, it now extends to the left - and including - point $-\lambda_2$, the standardised value of zero for period 2.[25]

Discretisation At this stage we define the approximation to the standardised actual WGD. This involves the partition of the support of $\tilde{w}_{ti}$ into J successive sub-intervals of equal length, denoted by $\mathcal{B}_j$, which we refer to as 'bins'

$$\mathcal{B}_j \equiv [\eta_{j-1}, \eta_j) \quad , \quad j = 1, \ldots, J \tag{4.41}$$

$$\eta \equiv \eta_j - \eta_{j-1} \quad \forall j \tag{4.42}$$

such that

$$\Pr\left(\tilde{w}_{ti} \in \mathcal{B}_\zeta\right) = 0 \quad , \quad \zeta \notin \{1, \ldots, J\} \tag{4.43}$$

$$\Pr\left(\tilde{w}_{ti} \in \bigcup_{j=1}^{J} \mathcal{B}_j\right) = 1 \tag{4.44}$$

and define the (ordered) discrete random variable y_{ti} whose value is equal to the value of the index of the bin that contains $\tilde{w}_{ti}$:

$$y_{ti} \equiv \{j,\ j \in \{1, \ldots, J\} : \tilde{w}_{ti} \in \mathcal{B}_j\} \tag{4.45}$$

$$y_{ti} \sim f_{yt}(\cdot) \tag{4.46}$$

[25] The 'standardization' procedure described here is a convenient tool to develop the semi-parametric estimator that is studied in this thesis. Its sole purpose is to alter the location of the actual WGDs so that the notional WGDs corresponding to their location-standardized versions are identical across periods, while, at the same time, *preserving the relative position of the notional distribution and the distributions of the rigidity bounds*, i.e. the point zero, in the case of DNWR (degenerate distribution), and the distribution of anticipated inflation rates, in the case of DRWR. In order to satisfy this condition, the (distributions of) the rigidity bounds must be shifted by the same distance as the actual WGD. Then, as a result, the position of the distortions introduced to the shape of the unstandardized actual WGD due to the presence of the two types of rigidity relative to the position of the unstandardized notional WGD, will be the same as the position of the distortions to the standardised actual WGD relative to the position of the standardised notional WGD. We note that this approach has also been used by others in the literature who adopted the location-histogram approach, e.g. Kahn (1997) and Beissinger and Knoppik (2001). Alternatively, and as we did in Chapters 2 and 3, one could proceed directly to the 'discretisation' stage, defining year specific bins relative to the position of some location parameter of the notional distribution for each year.

where $f_{yt}(\cdot)$ is the PMF, or *probability histogram*, of y_{ti}:

$$f_{yt} : \mathbb{N}_+ \to \{P_{jt}\}_{j=1,\dots,J} \tag{4.47}$$

$$P_{jt} \equiv \Pr(y_{ti} = j) = \Pr(\tilde{w}_{ti} \in \mathcal{B}_j) = \int_{\eta_{j-1}}^{\eta_j} \tilde{f}_t(\tilde{w})\, d\tilde{w} \tag{4.48}$$

Equivalently, the discretisation could be represented by the vector of binary scalars

$$\mathbf{d}_{ti} \equiv [d_{tij}]_{j=1,\dots,J} \tag{4.49}$$

$$d_{tij} \equiv I_{(\tilde{w}_{ti} \in \mathcal{B}_j)} \tag{4.50}$$

which has all its elements taking the value of zero[26] except for the element whose position coincides with the position of the bin that contains $\tilde{w}_{ti}$. Then, by construction, d_{tij} is a Bernoulli random variable, with mean P_{jt}

$$d_{tij} \sim Bernoulli(P_{jt}) \tag{4.51}$$

and, accordingly, $\mathbf{d}_{ti}$ a Multinomial random vector

$$\mathbf{d}_{ti} \sim Multinomial(1, \mathbf{P}_t) \quad , \quad \mathbf{P}_t \equiv [P_{jt}]_{j=1,\dots,J} \tag{4.52}$$

We also define

$$P_j^N \equiv \Pr(\tilde{w}_{ti}^N \in \mathcal{B}_j) \tag{4.53}$$

$$\pi_{jt} \equiv \Pr\left(\tilde{P}_{ti}^e \in \mathcal{B}_j\right) \tag{4.54}$$

to be the probabilities that the standardised notional wage growth rate, and standardised anticipated inflation rate of individual i for period t, respectively, fall in bin j. Finally, we introduce the following notation to indicate the value of index j for 'special' bins that will be needed for the exposition of the material that follows:

$$J_t^0 \equiv \{j,\ j \in \{1,\dots,J\} : -\lambda_t \in \mathcal{B}_j\} \tag{4.55}$$

$$J_t^P \equiv \left\{j,\ j \in \{1,\dots,J\} : E\tilde{P}_{ti}^e \in \mathcal{B}_j\right\} \tag{4.56}$$

$$\bar{J}_t^P \equiv \left\{j,\ j \in \{1,\dots,J\} : \max\left(\tilde{P}_{ti}^e\right) \in \mathcal{B}_j\right\} \tag{4.57}$$

$$\underline{J}_t^P \equiv \left\{j,\ j \in \{1,\dots,J\} : \min\left(\tilde{P}_{ti}^e\right) \in \mathcal{B}_j\right\} \tag{4.58}$$

i.e. J_t^0 is the index value of the bin that contains $-\lambda_t^N$ (the standardised value of zero), J_t^P the index value of the bin that contains the standardised value of mean anticipated inflation for period t, and $\underline{J}_t^P$ the index value of the leftmost and $\bar{J}_t^P$ the index value of the rightmost bins, respectively, that contain standardised values of anticipated inflation for period t.[27]

[26] The index function $I_{(\cdot)}$ is equal to 1, if the condition in the parentheses is satisfied, and 0 otherwise.

[27] Note that the bins that contain values of the standardised anticipated inflation distribution (i.e. values of j s.t. $\pi_{jt} > 0$) must have limits that satisfy the condition $\{\eta_j > \min(P_{ti}^e)\ \&\ \eta_{j-1} < \max(P_{ti}^e)\}$.

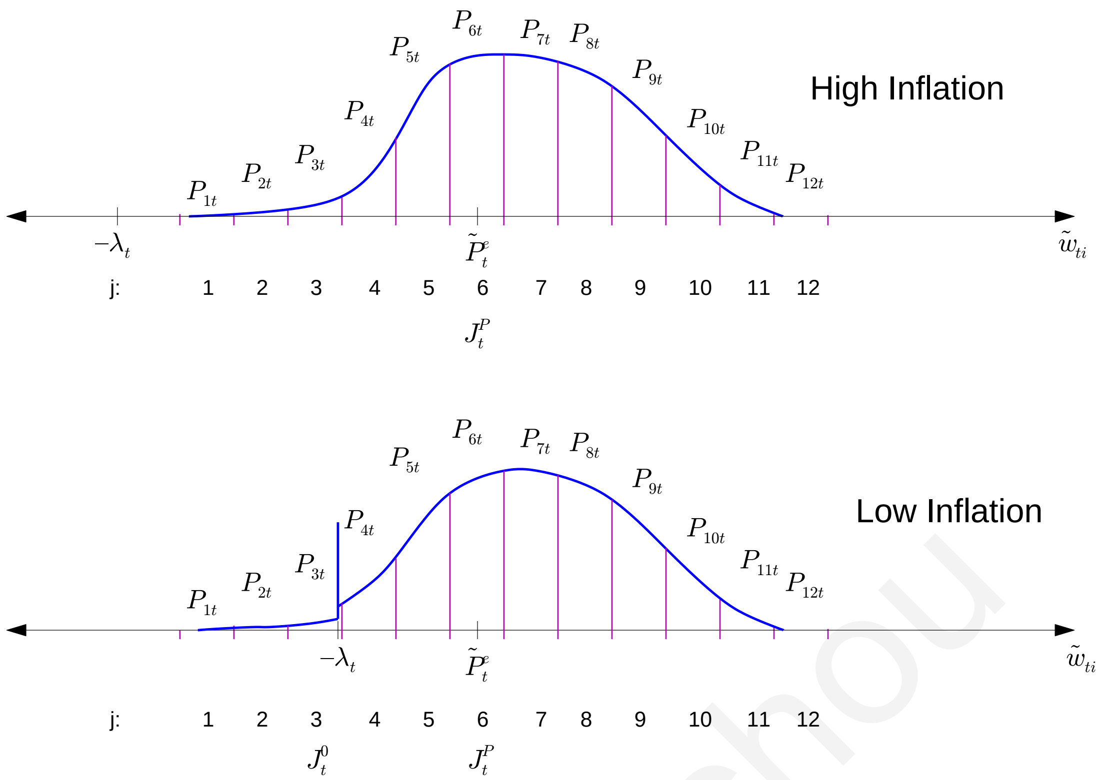


Figure 4.5: Discretisation.

An example of implementation of the discretisation stage is depicted in Figure 4.5. The solid line in the top diagram represents the PDF of a standardised WGD from a high inflation period, where DNWR had no effect on its shape. On the other hand, the solid line in the bottom diagram represents the PDF of a standardised WGD from a low inflation period, where DNWR had produced distortions that are visible for $\tilde{w}_{ti} \leq -\lambda_t$. Discretisation is implemented with $12(= J)$ bins. In both diagrams the bin that includes the standardised value of anticipated inflation has index value $j = 6$, therefore $J_t^P = 6$ in both cases. Furthermore, for the case depicted in the bottom diagram, the standardised value of zero, i.e. $-\lambda_t$, is located in bin $j = 3$ therefore, for this case, $J_t^0 = 3$.

Given the relationships that exist between the PDFs of the standardised notional and actual WGDs, and the standardised AID, a similar relationship can be derived for the PMFs of their respective discrete approximations. Using (4.37)-(4.40), we can write P_{jt}, defined in (4.48), as follows

$$\begin{aligned} P_{jt} &= \int_{\eta_{j-1}}^{\eta_j} \tilde{f}_t(\tilde{w})\, d\tilde{w} \\ &= \int_{\eta_{j-1}}^{\eta_j} \left[p^f \tilde{f}^N(\tilde{w}) + p^n \tilde{f}_t^n(\tilde{w}) + p^r \tilde{f}_t^r(\tilde{w}) \right] d\tilde{w} \\ &= p^f P_j^N + p_{jt}^n P_{jt}^n + p^r P_{jt}^r \end{aligned} \tag{4.59}$$

where

$$P_j^N = \int_{\eta_{j-1}}^{\eta_j} \tilde{f}^N(\tilde{w})\, d\tilde{w} \tag{4.60}$$

$$P_{jt}^{n} = \int_{\eta_{j-1}}^{\eta_j} \tilde{f}_t^n(\tilde{w})\, d\tilde{w} = \begin{cases} P_j^N & , \quad j > J_t^0 \\ P_j^N - \rho^n P_j^N & , \quad j < J_t^0 \\ P_j^N + \rho^n \sum_{\zeta<j} P_\zeta^N & , \quad j = J_t^0 \end{cases} \tag{4.61}$$

and

$$P_{jt}^{r} = \int_{\eta_{j-1}}^{\eta_j} \tilde{f}_t^r(\tilde{w})\, d\tilde{w} \tag{4.62}$$

$$= P_j^N - P_j^N \left[1 - \tilde{G}_t(\eta_j)\right] \rho^r + \left(\sum_{\zeta<j} P_\zeta^N\right) \left[\tilde{G}_t(\eta_j) - \tilde{G}_t(\eta_{j-1})\right] \rho^r \tag{4.63}$$

$$= P_j^N - P_j^N \left(\sum_{\xi>j} \pi_{\xi t}\right) \rho^r + \left(\sum_{\zeta<j} P_\zeta^N\right) \pi_{jt} \rho^r \tag{4.64}$$

$$= \begin{cases} P_j^N & , \quad j > \bar{J}_t^P \\ P_j^N - \rho_j^r P_j^N & , \quad j < \underline{J}_t^P \\ P_j^N - P_j^N \left(\sum_{\xi>j} \pi_{\xi t}\right) \rho^r + \left(\sum_{\zeta<j} P_\zeta^N\right) \pi_{jt} \rho^r & , \quad \underline{J}_t^P \leq j \leq \bar{J}_t^P \end{cases} \tag{4.65}$$

Then, combining, we can write:

$$\begin{aligned} P_{jt} = P_j^N - & \\ & - \underbrace{p^n \rho^n \; P_j^N \; I_{\left(j<J_t^0\right)}}_{DNWR\ loss} + \underbrace{p^n \rho^n \left(\sum_{\zeta<j} P_\zeta^N\right) \; I_{\left(j=J_t^0\right)}}_{DNWR\ gain} \\ & - \underbrace{p^r \rho^r \; P_j^N \left(\sum_{\xi>j} \pi_{\xi t}\right)}_{DRWR\ loss} + \underbrace{p^r \rho^r \left(\sum_{\zeta<j} P_\zeta^N\right) \; \pi_{jt}}_{DRWR\ loss} \end{aligned} \tag{4.66}$$

Construction of the likelihood function Given that $\dot{w}_{ti}$ are independent across i and t, the same must be true for $\tilde{w}_{ti}$ and y_{ti}. It then follows that the likelihood function for the sample of y_{ti}'s is given by

$$L_2 = \prod_{t=1}^{T} \prod_{i=1}^{n_t} \prod_{j=1}^{J} P_{jt}^{I_{(y_{ti}=j)}} \tag{4.67}$$

and the corresponding log-likelihood function is by

$$lnL_2 = \sum_{t=1}^{T} \sum_{i=1}^{n_t} \sum_{j=1}^{J} I_{(y_{ti}=j)} lnP_{jt} \tag{4.68}$$

At this point we note that the problem of maximising this log-likelihood is not tractable because the allocation of the probability mass of the anticipated inflation distribution (AID) across the bins that contain values of anticipated inflation may be different across years. This allocation is determined by the position of the AID relative to the notional WGD, since by assumption, the shape of the AID distribution is the same across years and only its location is allowed to vary. Therefore, in general, $\pi_{jt} \equiv \Pr\left(\tilde{P}^e_{ti} \in \mathcal{B}_j\right)$ can be expected to vary across t, and given that within each t the collection of probabilities $\{\pi_{jt}\}_{j=1,\ldots,J}$ and ρ^{rr} always appear multiplied to each other in the log-likelihood function above, this means that they cannot be identified separately.

The variation in the relative position of the notional WGD and AID would not pose a problem if the location of the standardised value of mean anticipated inflation ($E\tilde{P}^e_{ti}$) relative to the endpoints of the bin that contained it (i.e. the bin indexed by J^P_t) were fixed across t, i.e.[28]

$$E\tilde{P}^e_{ti} - \eta_{J^P_t - 1} = c \quad \Leftrightarrow \eta_{J^P_t} - E\tilde{P}^e_{ti} = \eta - c \tag{4.69}$$

Under this condition, the probability mass of AID that would fall in the bin that contained the standardised mean anticipated inflation would be the same across periods

$$\pi_{jt} = \Pr\left(\tilde{P}^e_{ti} \in \mathcal{B}_j\right) = \pi_0 \quad , \quad j = J^P_t \tag{4.70}$$

and the same would also be true for all bins that contained values of anticipated inflation, i.e.

$$\pi_{jt} = \Pr\left(\tilde{P}^e_{ti} \in \mathcal{B}_j\right) = \pi_q \quad , \quad q \equiv j - J^P_t \quad \& \quad j \in \left\{\underline{J}^P_t, \ldots, \bar{J}^P_t\right\} \tag{4.71}$$

where the index q gives the *relative position* of bin j in the support of the standardised AID.

Equation (4.64) would then become

$$P^r_{jt} = P^N_j - P^N_j \left(\sum_{\xi > j} \pi_{\xi - J^P_t}\right) \rho^r + \left(\sum_{\zeta < j} P^N_\zeta\right) \pi_{j - J^P_t} \rho^r \tag{4.72}$$

and equation (4.59) would simplify to the following[29]

$$P_{jt} = P^N_j + D^n_{jt} + D^r_{jt} \tag{4.73}$$

[28]Given that the bin width is the same for all bins, this is equivalent to having the difference between any two values taken by the standardised value of mean anticipated inflation across periods to be a multiple of the bin width: $|E\dot{P}^e_{ti} - E\dot{P}^e_{\tau i}| = \nu \cdot \eta \ , \quad \nu \in \mathbb{N}_+, \quad t \neq \tau$.

[29]We note that this result concerning the bin heights ($= P_{jt}$) coincides with the result in (2.7) in Chapter 2, except that there we also allow for distortions due to menu costs. The additive nature of the distortions in both expressions is the result of, explicitly or implicitly, assuming that the population members are independently assigned to the different wage-setting regimes. However, beyond this similarity, the rigidity mechanism described in this chapter is different from the one implicitly assumed in Chapters 2 and 3, and so are the functional forms of D^n_{jt} and D^r_{jt} in the two cases. In particular, in this chapter the size of the distortions is proportional to the bin heights, and therefore, for each type of rigidity, this is fully described by the proportionality parameters ρ^{nn} and ρ^{rr}, given knowledge of the notional bin height. Consequently, these are the relevant rigidity parameters to be estimated for this case, given that we also estimate the notional bin heights. On the other hand, in Chapters 2 and 3 the size of the distortions does not poses this proportionality feature, and therefore estimating ρ^{nn} and ρ^{rr} for these cases is not relevant. In those cases the distortions are, instead, allowed to take any value given certain restrictions that do not relate to the notional bin heights, and therefore the size of the distortion relevant to each bin has to be estimated separately.

where

$$D^n_{jt} = \underbrace{-\overbrace{p^n\rho^n}^{\rho^{nn}} P^N_j \, I_{\left(j<J^0_t\right)}}_{DNWR\ loss} + \underbrace{\overbrace{p^n\rho^n}^{\rho^{nn}} \left(\sum_{\zeta<j} P^N_\zeta\right) I_{\left(j=J^0_t\right)}}_{DNWR\ gain} \tag{4.74}$$

$$D^r_{jt} = \underbrace{-\overbrace{p^r\rho^r}^{\rho^{rr}} \; P^N_j \left(\sum_{\xi>j} \pi_{\xi-J^P_t}\right)}_{DRWR\ loss} + \underbrace{\overbrace{p^r\rho^r}^{\rho^{rr}} \left(\sum_{\zeta<j} P^N_\zeta\right) \pi_{j-J^P_t}}_{DRWR\ gain} \tag{4.75}$$

The log-likelihood function would then be given by

$$lnL_2 = \sum_{t=1}^{T} \sum_{i=1}^{n_t} \sum_{j=1}^{J} I_{(y_{ti}=j)} ln\left(P^N_j + D^n_{jt} + D^r_{jt}\right) \tag{4.76}$$

For this case, the unknown parameters to be estimated through the maximisation of the log-likelihood function would then be

$$\begin{aligned} \rho^{nn} &: \text{ Overall DNWR effect} \\ \rho^{rr} &: \text{ Overall DRWR effect} \\ \left\{P^N_j\right\}_{j=1,\ldots,J} &: \text{ probability-histogram of standardised notional distribution} \\ \left\{\pi_q\right\}_{q\in\mathcal{Q}} &: \text{ probability-histogram of standardised AID} \end{aligned} \tag{4.77}$$

where

$$\mathcal{Q} \equiv \left\{q = j - J^P_t : j \in \left\{\underline{J}^P_t, \ldots, \bar{J}^P_t\right\}\right\} \tag{4.78}$$

is defined to be the set of values taken by the index q.[30]

4.4.2 Identification

In this part we discuss sufficient conditions for the identification of the unknown parameters of the model presented in Section 4.2. First we consider the case where only DNWR is present, then the case where only DRWR is present, and conclude with the case where both types of rigidity are present.

DNWR

The data from a particular period, indexed by t,

$$\left\{\dot{w}_{ti}\right\}_{i=1,\ldots,n_t} \tag{4.79}$$

[30]For example, if the number of bins covered by the support of the AID is equal to three (in all periods), then $q \in \mathcal{Q} = \{-1, 0, 1\}$, where $q = 0$ corresponds to the bin that contains mean anticipated inflation in the given period (i.e. $j = J^P_t$) , $q = -1$ to the bin located one position to its left, and $q = 1$ to the bin located one position to its right.

allows us to estimate the set of probabilities

$$\{P_{jt}\}_{j=1,\dots,J} \tag{4.80}$$

i.e. the height of the bins of the probability histogram corresponding to the standardised actual WGD[31] (the 'actual' probability histogram) for that period. We recall from (4.61) that, in the presence of DNWR, the quantities above relate to the parameter of interest ρ^n and the notional bin heights $\left\{P_j^N\right\}_{j=1,\dots,J}$, which are ancillary parameters, in the following way[32]

$$P_{jt} = \begin{cases} P_j^N & , \quad j > J_t^0 \\ (1-\rho^n)\, P_j^N & , \quad j < J_t^0 \\ P_j^N + \rho^n \sum_{\zeta<j}^N P_\zeta^N & , \quad j = J_t^0 \end{cases} \tag{4.81}$$

The following theorem provides sufficient conditions for the identification of ρ^n:

Theorem 1 *Conditions on the data on actual wage growth rates*

$$\{\dot{w}_{ti}\}_{t=1,\dots,T,\; i=1,\dots,n_t} \tag{4.82}$$

that are sufficient for the identification of the rigidity parameter ρ^n in the semiparametric model described above, are the following:

1. *$T \geq 2$, and*
2. *for at least one pair of periods $t_1, t_2 \in \{1, \dots, T\}$, where $t_1 \neq t_2$, the following are satisfied:*
   - *(a) $1 < J_{t_1}^0 \leq J$, i.e. the 'actual' probability histogram in $t = t_1$ is distorted by DNWR,*
   - *(b) $J_{t_2}^0 < J_{t_1}^0$, i.e. at least one of the bins distorted in $t = t_1$ is not distorted in $t = t_2$.*

Proof. Without loss of generality we consider the case where

$$T = 2 \tag{4.83}$$

$$1 < J_2^0 < J_1^0 \leq J \tag{4.84}$$

i.e. the probability histograms underlying the data from the two periods are both distorted by DNWR, since both $J_1^0 > 1$ and $J_2^0 > 1$, but not all bins whose height is distorted in $t = 1$ are also affected in $t = 2$. An example that adheres to this case is depicted in Figure 4.6, where the solid line represents the PDF of the standardised notional WGD ($\tilde{f}^N$), the wider bins correspond to the 'actual' probability histogram of for period 1 ($P_{1j},\ j = 1, \dots, J$), and the narrower bins to the 'actual' probability histogram for period 2 ($P_{2j},\ j = 1, \dots, J$).

[31] That is, the PMF of y_{ti}.

[32] Here we assume that $p^n = 1$. In the case where $p^n \in (0,1)$ and $p^f \in (0,1)$ (and $p^r = 0$), then the results derived here would be valid for $\rho^{nn} \equiv p^n \rho^n$ rather than ρ^n.

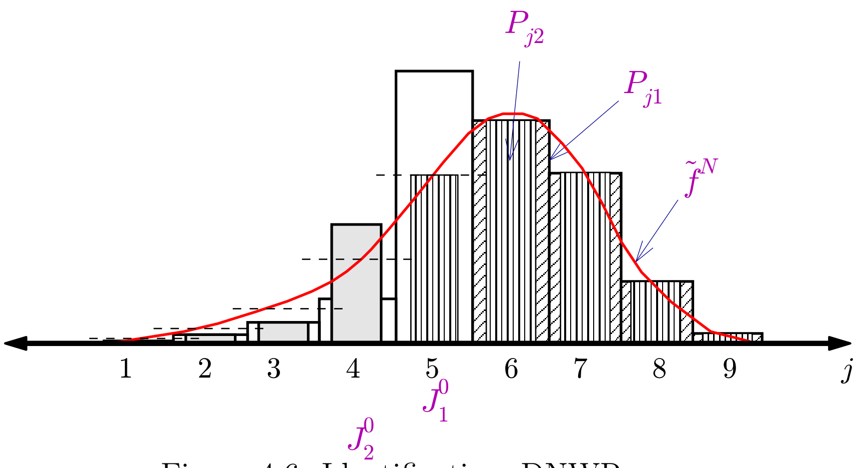


Figure 4.6: Identification, DNWR case.

In this example we have $J = 9$. In period 1 the bin that contains the standardised value of zero has index value $j = 5$ $(= J_1^0)$, and for period 2 the corresponding bin has index value $j = 4$ $(= J_2^0)$. For ease of reference we mark the bins - in both histograms - that are not distorted by rigidity[33] with stripes. We also use a dashed horizontal line to indicate the height of the corresponding bin in the notional probability histogram for those values of j where at least one of the bins is distorted.

For the general case defined by (4.83) and (4.84), the available data

$$\{\dot{w}_{ti}\}_{t=1,2,\ i=1,\dots,n_t} \tag{4.85}$$

allows us to obtain the collection of estimates

$$\left\{\hat{P}_{jt}\right\}_{j=1,\dots,J} \quad , \quad t = 1, 2 \tag{4.86}$$

where, in period $t = 1$, the estimated quantities satisfy

$$P_{j1} = \begin{cases} P_j^N & , \quad j = J_1^0 + 1, \dots, J \quad (\text{if } J_1^0 < J) \\ (1 - \rho^n)\, P_j^N & , \quad j = 1, \dots, J_1^0 - 1 \\ P_j^N + \rho^n \sum_{\zeta < j} P_\zeta^N & , \quad j = J_1^0 \end{cases} \tag{4.87}$$

and, in period $t = 2$

$$P_{j2} = \begin{cases} P_j^N & , \quad j = J_2^0 + 1, \dots, J_1^0, \dots, J \\ (1 - \rho^n)\, P_j^N & , \quad j = 1, \dots, J_2^0 - 1 \\ P_j^N + \rho^n \sum_{\zeta < j} P_\zeta^N & , \quad j = J_2^0 \end{cases} \tag{4.88}$$

Thus, for the example depicted in Figure 4.6, we see that the wider bins - that correspond to the period 1 histogram - have equal height to the corresponding notional bins (and are, therefore, marked by stripes) for $j \geq 5$ $(= J_1^0)$, the bin indexed by $j = 5$ is taller than the corresponding notional bin (and not marked by stripes), and all bins indexed by $j < 5$ are shorter than the corresponding notional bins (and also not marked by stripes). A similar

[33] Whose height is, therefore, the same as the height of the corresponding bins of the 'notional' probability histogram; this, by assumption, is the same for both periods.

pattern applies to the period 2 (narrower) bins, except that in that case the three different type of bins discussed above - equal, taller, shorter than the corresponding notional bins - are defined with respect to the bin indexed by $j = 4 (= J_2^0)$.

The probabilities from the two periods which are associated with the bin that contains the standardised value of point zero in period 1, i.e. the bin indexed by $j = J_1^0$, satisfy

$$P_{j1} = P_j^N + \rho^n \sum_{\zeta<j} P_\zeta^N \tag{4.89}$$

$$P_{j2} = P_j^N \tag{4.90}$$

Combining these two, we can write

$$\rho^n = \frac{P_{j1} - P_{j2}}{\sum_{\zeta<j} P_\zeta^N} \quad , \quad j = J_1^0 \tag{4.91}$$

Then, using the result in the top row on the RHS of (4.88), we can write the denominator as a function of the quantities identified by the data from period $t = 2$;

$$\sum_{\zeta<J_1^0} P_\zeta^N = 1 - \sum_{\zeta\geq J_1^0} P_\zeta^N = 1 - \sum_{\zeta\geq J_1^0} P_{\zeta 2} \tag{4.92}$$

and, therefore, substituting back into (4.91), we could express the rigidity parameter as a function of quantities that are estimable

$$\rho^n = \frac{P_{j1} - P_{j2}}{1 - \sum_{\zeta\geq j} P_{\zeta 2}} \quad , \quad j = J_1^0 \tag{4.93}$$

■

The above result is self explanatory: the numerator gives the magnitude of the probability mass that, due to DNWR, is missing, in total, from the bins that lie to the left of the bin that contains the standardised value of zero in period 1, and correspond to the area of the un-standardised notional WGD that lies below zero.[34] On the other hand, the denominator gives the probability mass that would be allocated to those bins in the absence of DNWR. Thus their ratio gives the probability of the event a negative notional wage growth rate not being realised conditional on the event the notional growth rate being negative, which is the formal definition of ρ^n.[35] This ratio is often referred to in this literature as the ‘sweep-up ratio’, which, in the case of proportional DNWR, is equal to ρ^n.

We conclude this part by making three additional points:

First, we note from (4.88) that the data from period $t = 2$ allow the identification of $\{P_j^N\}_{j=J_2^0+1,\ldots,J_1^0,\ldots,J}$. From (4.87) we also note that

$$P_{j1} = \ (1 - \rho^n)\, P_j^N \quad , \quad j = 1, \ldots, J_1^0 - 1 \tag{4.94}$$

Therefore we can substitute expression (4.93) for ρ^n in (4.94) and solve for P_j^N to write the

[34] More precisely, below $\eta_{J_1^0-1} + \lambda_1$.

[35] See footnote 11.

remaining probabilities of the 'notional' probability histogram as a function of identifiable quantities.

Secondly, we note that an alternative identification strategy is available in the special case where, in addition to the conditions in (4.84), it is also true that $J_1^0 - J_2^0 > 1$, i.e. the difference in the value of the index of the bins that contain the standardised values of zero in the histograms from the two periods is greater than one. In that case, for those bins indexed by $j = J_2^0 + 1, \ldots, J_1^0 - 1$, i.e. the bins 'in between' the bins indexed by $j = J_2^0$ and $j = J_1^0$, it is true that

$$P_{j1} = (1 - \rho^n)\, P_j^N \tag{4.95}$$

$$P_{j2} = P_j^N \tag{4.96}$$

therefore we could write

$$\rho^n = \frac{P_{j2} - P_{j1}}{P_{j2}} \quad , \quad J_2^0 < j < J_1^0 \tag{4.97}$$

where, as in (4.93), the quantities on the RHS of this expression are all identified by the available data.

Thirdly, we note that both formulas in (4.93) and (4.97) are also valid in the case where $J_2^0 \leq 1$, i.e. the location of the second distribution is in the positive orthant and far from point zero such that its probability histogram cannot be distorted by the presence of DNWR. In that case we have

$$P_{j2} = P_j^N \quad , \quad j = 1, \ldots, J \tag{4.98}$$

DRWR

In this case, the data from a particular period, indexed by t,

$$\{\dot{w}_{ti}\}_{i=1,\ldots,n_t} \tag{4.99}$$

allows us to estimate the 'actual' probability histogram bin heights[36]

$$\{P_{jt}\}_{j=1,\ldots,J} \tag{4.100}$$

that satisfy[37]

$$P_{jt} = \begin{cases} P_j^N & , \quad j > \bar{J}_t^P \\ (1 - \rho^r)\, P_j^N & , \quad j < \underline{J}_t^P \\ P_j^N - P_j^N \rho^r \left(\sum_{\xi > j} \pi_{\xi - \underline{J}_t^P}\right) + \left(\sum_{\zeta < j} P_\zeta^N\right) \rho^r \pi_{j - \underline{J}_t^P} & , \quad \underline{J}_t^P \leq j \leq \bar{J}_t^P \end{cases} \tag{4.101}$$

[36]Here we assume that $p^r = 1$. In the case where $p^r \in (0,1)$ and $p^f \in (0,1)$ (and $p^n = 0$), then the results derived here would be valid for $\rho^{rr} \equiv p^r \rho^r$ rather than ρ^r. We also assume that the location of AID is 'standardised' in the way discussed at the end of Section 4.4.1.

[37]A simplified version of the relationship given in (4.73)-(4.75), where $p^r = 1$ (thus $p^f = p^n = 0$).

The following theorem provides sufficient conditions for the identification of ρ^r:

Theorem 2 *Conditions on the data on actual wage growth rates*

$$\{\dot{w}_{ti}\}_{t=1,\ldots,T,\ i=1,\ldots,n_t} \tag{4.102}$$

that are sufficient for the identification of the rigidity parameter ρ^r in the semiparametric model described above, are the following:

1. *$T \geq 2$, and*
2. *for at least one pair of periods $t_1, t_2 \in \{1, \ldots, T\}$, where $t_1 \neq t_2$, the following are satisfied:*
   - *(a) $1 \leq \underline{J}_{t_1}^P < \bar{J}_{t_1}^P \leq J$, i.e. the support of the AID in $t = t_1$ is a subset of the support of the notional WGD in that period,*
   - *(b) $\bar{J}_{t_2}^P < \underline{J}_{t_1}^P$, i.e. the support of the AID in $t = t_2$ does not overlap with the support of the AID in $t = t_1$ and lies to its left, therefore the bins that contain the support of AID in $t = t_1$ are not distorted in $t = t_2$ by DRWR.*

Proof. Without loss of generality we consider the case where

$$T = 2 \tag{4.103}$$

$$1 \leq \underline{J}_{t_1}^P < \bar{J}_{t_1}^P \leq J \tag{4.104}$$

$$1 \leq \bar{J}_{t_2}^P < \underline{J}_{t_1}^P \tag{4.105}$$

i.e. the underlying probability histograms for the data from the two periods are both distorted by DRWR, but the bins that contain the support of the AID in period $t = 1$ are not distorted in period $t = 2$. An example that adheres to the above assumptions is depicted in Figure 4.7,[38] which shows the PDF of the standardised notional WGD, as well as the 'actual' probability histograms for periods 1 and 2, where $J_1^P = 6$ and $J_2^P = 2$. Furthermore, the support of the (standardised) AID covers three bins such that, for period 1, $\underline{J}_1^P = 5$ and $\bar{J}_1^P = 7$, and, for period 2, $\underline{J}_2^P = 2$ and $\bar{J}_2^P = 4$.

For the general case described by (4.103)-(4.105), the available data

$$\{\dot{w}_{ti}\}_{t=1,2,\ i=1,\ldots,n_t} \tag{4.106}$$

allows us to obtain the collection of estimates

$$\left\{\hat{P}_{jt}\right\}_{j=1,\ldots,J} \quad , \quad t = 1, 2 \tag{4.107}$$

where in period 1, the estimated quantities satisfy

$$P_{j1} = \begin{cases} P_j^N & , \quad j = \bar{J}_1^P + 1, \ldots, J \\ (1-\rho^r)\, P_j^N & , \quad j = 1, \ldots, \underline{J}_1^P - 1 \\ \left.\begin{array}{l} P_j^N - P_j^N \rho^r \left(\sum_{\xi=j+1}^{\bar{J}_1^P} \pi_{\xi - J_1^P}\right) + \\ + \left(\sum_{\zeta<j} P_\zeta^N\right) \rho^r \pi_{j-J_1^P} \end{array}\right\} & , \quad j = \underline{J}_1^P, \ldots, \bar{J}_1^P \end{cases} \tag{4.108}$$

[38]The format of the design is the same to that of Figure 4.6, which depicted the DNWR example.

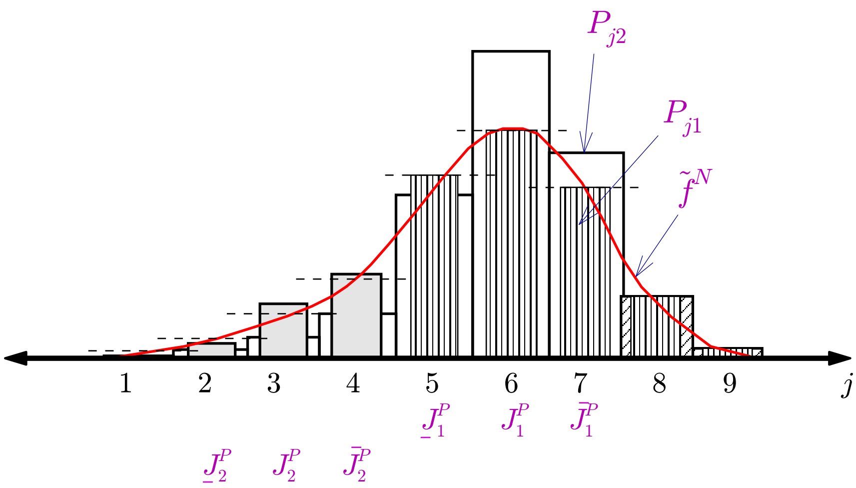


Figure 4.7: Identification, DRWR case.

and, in period 1

$$P_{j2} = \begin{cases} P_j^N & , \quad j = \bar{J}_2^P + 1, \dots, \underline{J}_1^P, \dots, J \\ (1 - \rho^r)\, P_j^N & , \quad j = 1, \dots, \underline{J}_2^P - 1 \\ \left. \begin{array}{l} P_j^N - P_j^N \rho^r \left(\sum_{\xi=j+1}^{\bar{J}_2^P} \pi_{\xi - J_2^P} \right) + \\ + \left(\sum_{\zeta<j} P_\zeta^N \right) \rho^r \pi_{j - J_2^P} \end{array} \right\} & , \quad j = \underline{J}_2^P, \dots, \bar{J}_2^P \end{cases} \tag{4.109}$$

In Figure 4.7 we see that for period 1 (wider bins), the bins indexed by $j > 7 (= \bar{J}_1^P)$ have equal height to the notional bins (see first row of (4.108)), the bins indexed by $j < 5 (= \underline{J}_1^P)$ exhibit a deficit of probability mass relative to the notional bins (see second row of (4.108)), while the bins indexed by $5 \leq j \leq 7$ - that contain the support of the AID in period 1 - exhibit either deficit or surplus (see third row of (4.108)): the leftmost bin ($j = 5$) exhibits a deficit, while the other two ($j = 6, 7$) a surplus. The same pattern can be seen in the histogram for the period 2 except that, for that case, the three types of bins discussed above have different location: the bins that have equal height to the notional bins are indexed by $j > 4 (= \bar{J}_2^P)$, the one that exhibits deficit is indexed by $j = 1 (< \underline{J}_2^P = 2)$, and, finally, the bins indexed by $2 \leq j \leq 4$ - that contain the support of AID for that period - exhibit either deficit or surplus.

From (4.108) and (4.109) it then follows that, for the bin that contains the maximum value of the support of the AID in period $t = 1$, i.e. for $j = \bar{J}_1^P$, the associated probabilities from the two periods satisfy

$$P_{j1} = P_j^N - P_j^N \underbrace{\left(\sum_{\xi=j+1}^{\bar{J}_1^P} \pi_{\xi - J_1^P} \right)}_{=0} \rho^r + \left(\sum_{\zeta<j} P_\zeta^N \right) \rho^r \pi_{j-J_1^P} \tag{4.110}$$

$$= P_j^N + \left(\sum_{\zeta<j} P_\zeta^N \right) \rho^r \pi_{j-J_1^P} \tag{4.111}$$

$$P_{j2} = P_j^N \tag{4.112}$$

Furthermore, for period $t = 2$ and $j = \bar{J}_1^P$, it is also true that (see Figure 4.7, for $j = 7$):

$$1 - \sum_{\zeta\geq j} P_{\zeta 2} = 1 - \sum_{\zeta\geq j} P_\zeta^N = \sum_{\zeta<j} P_\zeta^N \tag{4.113}$$

Combining the above, we can write

$$\rho^r \pi_{j-J_1^P} = \frac{P_{j1} - P_{j2}}{1 - \sum_{\zeta\geq j} P_{\zeta 2}} \quad , \quad j = \bar{J}_1^P \tag{4.114}$$

In the same way, using (4.108) and (4.109), we can also write, for $j = \bar{J}_1^P - 1$ (i.e. for the bin located one position to the left of the bin indexed by $j = \bar{J}_1^P$ - see Figure 4.7, for $j = 6$) the following;

$$P_{j1} = P_j^N - P_j^N \left(\rho^r \pi_{\bar{J}_1^P - J_1^P} \right) + \left(\sum_{\zeta<j} P_\zeta^N \right) \rho^r \pi_{j-J_1^P} \tag{4.115}$$

$$P_{j2} = P_j^N \tag{4.116}$$

$$\sum_{\zeta<j} P_\zeta^N = 1 - \sum_{\zeta\geq j} P_{\zeta 2} \tag{4.117}$$

and therefore, combining (4.115)-(4.117), we can write

$$\rho^r \pi_{j-J_1^P} = \frac{P_{j1} - P_{j2} + P_{j2} \left(\rho^r \pi_{\bar{J}_1^P - J_1^P} \right)}{1 - \sum_{\zeta\geq j} P_{\zeta 2}} \quad , \quad j = \bar{J}_1^P - 1 \tag{4.118}$$

At this point it is clear that expressions (4.114) and (4.118) share the same structure, being special cases of the following expression

$$\rho^r \pi_{j-J_1^P} = \frac{P_{j1} - P_{j2} + P_{j2} \sum_{\xi=j+1}^{\bar{J}_1^P} \left(\rho^r \pi_{\xi - J_1^P} \right)}{1 - \sum_{\zeta\geq j} P_{\zeta 2}} \tag{4.119}$$

In fact, following the same line of thought as to the one used to derive (4.114) and (4.118), we can show that the following are true for all $j = \underline{J}_1^P, \ldots, \bar{J}_1^P$

$$P_{j1} = P_j^N - P_j^N \sum_{\xi=j+1}^{\bar{J}_1^P} \left(\rho^r \pi_{\xi - \underline{J}_1^P}\right) + \left(\sum_{\zeta<j} P_\zeta^N\right) \rho^r \pi_{j-\underline{J}_1^P} \tag{4.120}$$

$$P_{j2} = P_j^N \tag{4.121}$$

$$\sum_{\zeta<j} P_\zeta^N = 1 - \sum_{\zeta \geq j} P_{\zeta 2} \tag{4.122}$$

This would then lead us to derive expression (4.119), which is therefore true for all these values of j, that is

$$\rho^r \pi_{j-\underline{J}_1^P} = \frac{P_{j1} - P_{j2} + P_{j2} \sum_{\xi=j+1}^{\bar{J}_1^P} \left(\rho^r \pi_{\xi-\underline{J}_1^P}\right)}{1 - \sum_{\zeta \geq j} P_{\zeta 2}} \quad , \quad j = \underline{J}_1^P, \ldots, \bar{J}_1^P \tag{4.123}$$

In this expression we note that its RHS includes several probabilities from the two 'actual' probability histograms, i.e. a subset of the probabilities in (4.107), that can be identified from the available data, as well as the sum of terms $\rho^r \pi_{\xi-\underline{J}_1^P}$ for $\xi = j+1, \ldots, \bar{J}_1^P$.

This sum is equal to zero for $j = \bar{J}_1^P$, and as we can see from (4.114), the term $\rho^r \pi_{\bar{J}_1^P - \underline{J}_1^P}$ can be written as a function of only identifiable quantities.

For the bin immediately to its left, indexed by $j = \bar{J}_1^P - 1$, this sum will be equal to $\rho^r \pi_{\bar{J}_1^P - \underline{J}_1^P}$, therefore we can also write $\rho^r \pi_{j-\underline{J}_1^P}$ for $j = \bar{J}_1^P - 1$ as a function of only identifiable quantities.

Repeating this sequence of calculations, moving one bin at a time to the left, one can see that it is possible to express the RHS of (4.123) only as a function of identifiable quantities.

Then, using that

$$\sum_{\xi=\underline{J}_1^P}^{\bar{J}_1^P} \pi_{\xi-\underline{J}_1^P} = 1 \tag{4.124}$$

we can obtain ρ^r by summing-up the quantities on the RHS of (4.123) for $j = \underline{J}_1^P, \ldots, \bar{J}_1^P$:

$$\rho^r = \sum_{\xi=\underline{J}_1^P}^{\bar{J}_1^P} \rho^r \pi_{\xi-\underline{J}_1^P} \tag{4.125}$$

■

Having been able to write ρ^r as a function of identifiable quantities, it is then possible to recover the AID probabilities, $\{\pi_q\}_{q\in\mathcal{Q}}$, and the remaining notional probabilities (for $j < \underline{J}_1^P$). First we note that we can substitute the RHS of (4.125) into (4.123) and solve for $\pi_{j-\underline{J}_1^P}$, therefore in this way the probabilities in the 'AID' probability histogram can also be identified. Furthermore, we note from (4.109) that the data from period $t = 2$ allow the identification of $\left\{P_j^N\right\}_{j=\bar{J}_2^P+1,\ldots,\underline{J}_1^P,\ldots,J}$. From (4.108) we also note that

$$P_{j1} = \left(1 - \rho^r\right) P_j^N \quad , \quad j = 1, \ldots, \underline{J}_1^P - 1 \tag{4.126}$$

Therefore we can substitute expression (4.125) for ρ^r in (4.126) and solve for P_j^N to write the remaining probabilities of the 'notional' probability histogram as a function of identifiable quantities.

We conclude this part by making two additional observations. Firstly, we observe that in the special case where $\bar{J}_1^P = \underline{J}_1^P = J_1^P$, i.e. the support of the AID in $t = 1$ is confined within the width of a single bin, then for $j = J_1^P$ it is true that

$$\pi_{j-J_1^P} = \pi_0 = 1 \tag{4.127}$$

$$\sum_{\xi=j+1}^{\bar{J}_1^P} \pi_{\xi-J_1^P} = 0 \tag{4.128}$$

therefore expression (4.123) simplifies to the following

$$\rho^r \pi_{j-J_1^P} = \rho^r = \frac{P_{j1} - P_{j2}}{1 - \sum_{\zeta \geq j} P_{\zeta 2}} \quad , \quad j = J_1^P \tag{4.129}$$

This is similar to the result presented in expression (4.93) for the case of DNWR, where the support of the distribution of the rigidity bounds is degenerate at point zero and therefore also confined within the width of a single bin, the one indexed by $j = J_t^0$.

Secondly, in the special case where $\underline{J}_1^P - \bar{J}_2^P > 1$, an alternative identification strategy is also available. In that case we have, for $\bar{J}_2^P < j < \underline{J}_1^P$

$$P_{j1} = (1 - \rho^r)\, P_j^N \tag{4.130}$$

$$P_{j2} = P_j^N \tag{4.131}$$

and therefore we can write

$$\rho^r = \frac{P_{j2} - P_{j1}}{P_{j2}} \quad , \quad j = \bar{J}_2^P + 1, \ldots, \underline{J}_1^P - 1 \tag{4.132}$$

where the RHS includes only identifiable quantities.

DNWR and DRWR

In this case, the data from a particular period, indexed by t,

$$\{\dot{w}_{ti}\}_{i=1,\ldots,n_t} \tag{4.133}$$

allows us to estimate the bin heights of the 'actual' probability histogram

$$\{P_{jt}\}_{j=1,\ldots,J} \tag{4.134}$$

that satisfy

$$P_{jt} = P_j^N + D_{jt}^n + D_{jt}^r \tag{4.135}$$

where[39]

$$D^n_{jt} = -\underbrace{\rho^{nn}}_{=p^n\rho^n} P^N_j \; I\left(j < J^0_t\right) + \rho^{nn} \left(\sum_{\zeta<j} P^N_\zeta\right) \; I\left(j = J^0_t\right) \tag{4.136}$$

$$D^r_{jt} = -\underbrace{\rho^{rr}}_{=p^r\rho^r} \; P^N_j \left(\sum_{\xi>j} \pi_{\xi - J^P_t}\right) + \rho^{rr} \left(\sum_{\zeta<j} P^N_\zeta\right) \; \pi_{j-J^P_t} \tag{4.137}$$

The following theorem provides sufficient conditions for the identification of $\rho^{nn}$and ρ^{rr}:

Theorem 3 *Conditions on the data on actual wage growth rates*

$$\{\dot{w}_{ti}\}_{t=1,\dots,T,\; i=1,\dots,n_t} \tag{4.138}$$

that are sufficient for the identification of the total-rigidity parameters $\rho^{nn}$and ρ^{rr} in the semiparametric model described in Section 4.2, are the following:

1. $T \geq 2$,
2. *for at least one pair of periods $t_1, t_2 \in \{1, \dots, T\}$, where $t_1 \neq t_2$, the following are satisfied:*
   (a) $1 \leq \underline{J}^P_{t_1} < \bar{J}^P_{t_1} \leq J$ *and* $J^0_{t_1} < \underline{J}^P_{t_1}$, *i.e. the bins of the 'actual' probability histogram in $t = t_1$, that contain the support of AID are a subset of the support of the notional WGD in that period, and are only distorted by DRWR, while the remaining bins that lie to their left are possibly distorted by DNWR,*
   (b) $\bar{J}^P_{t_2} < \underline{J}^P_{t_1}$ *and* $J^0_{t_2} < \underline{J}^P_{t_1}$, *i.e. the bins that contain the support of AID in $t = t_1$ are not distorted in $t = t_2$ by either DRWR or DNWR, while any remaining bins that lie to their left are possibly distorted by DNWR, and*
3. *for at least one pair of periods $t_3, t_4 \in \{1, \dots, T\}$, where $t_3 \neq t_4$ and, possibly, $t_3, t_4 \in \{t_1, t_2\}$, the following are satisfied:*
   (a) $1 < J^0_{t_3} \leq J$, *i.e. the 'actual' probability histogram in $t = t_3$ is distorted by DNWR,*
   (b) $J^0_{t_3} < J^0_{t_4}$, *i.e. at least one of the bins distorted by DNWR in $t = t_3$ is not distorted in $t = t_4$ by this type of rigidity.*

Proof. Without loss of generality we assume that

$$T = 2 \tag{4.139}$$

$$1 \leq \underline{J}^P_{t_1} < \bar{J}^P_{t_1} \leq J \text{ and } 1 < J^0_{t_1} < \underline{J}^P_{t_1} \tag{4.140}$$

$$1 \leq \underline{J}^P_{t_2} \leq \bar{J}^P_{t_2} < \underline{J}^P_{t_1} \leq J \text{ and } 1 < J^0_{t_2} < \underline{J}^P_{t_1} \tag{4.141}$$

$$J^0_{t_2} \neq J^0_{t_1} \tag{4.142}$$

[39]See page 83. We also assume that the location of AID is 'standardised' in the way discussed at the end of Section 4.4.1.

i.e. the underlying probability histograms for the data from the two periods are distorted by both DNWR and DRWR, such that the bins that contain the support of the AID in period $t = 1$ are not distorted in period $t = 2$ by either of the two types of rigidity, and also, different sets of bins are distorted by DNWR in the two periods. Assumption (4.142) includes two sub-cases, namely

$$\text{Case 1}: \; J^0_{t_2} < J^0_{t_1} \tag{4.143}$$

$$\text{Case 2}: \; J^0_{t_1} < J^0_{t_2} \tag{4.144}$$

Here we will prove sufficiency of the above conditions for Case 1.

Without loss of generality, in addition to (4.139)-(4.142), we assume further that

$$J^0_{t_2} + 1 = \bar{J}^P_{t_2} = J^0_{t_1} = \underline{J}^P_{t_1} - 1 \tag{4.145}$$

Therefore, for Case 1 above, assumptions (4.139)-(4.142) and (4.145) imply the following order for the index values:[40]

$$1 \leq \underline{J}^P_{t_2} < J^0_{t_2} < \bar{J}^P_{t_2} \leq J^0_{t_1} < \underline{J}^P_{t_1} < \bar{J}^P_{t_1} \leq J \tag{4.146}$$

This case is depicted in Figure 4.8, where

	$\underline{J}^P_t$	J^P_t	$\bar{J}^P_t$	J^0_t
Period $t_1 = 1$*:*	5	6	7	4
Period $t_2 = 2$*:*	2	3	4	3

The available data

$$\{\dot{w}_{ti}\}_{t=1,2,\; i=1,\ldots,n_t} \tag{4.147}$$

allows us to obtain the collection of estimates

$$\left\{\hat{P}_{jt}\right\}_{j=1,\ldots,J} \quad , \quad t = 1, 2 \tag{4.148}$$

where in period 1, the estimated quantities satisfy

$$P_{j1} = \begin{cases} P^N_j & , \quad j = \bar{J}^P_1 + 1, \ldots, J \\[2ex] \left. \begin{array}{l} P^N_j - P^N_j \rho^{rr} \left(\sum_{\xi=j+1}^{\bar{J}^P_1} \pi_{\xi - J^P_1} \right) + \\ + \left(\sum_{\zeta<j} P^N_\zeta \right) \rho^{rr} \pi_{j - J^P_1} \end{array} \right\} & , \quad j = \underline{J}^P_1, \ldots, \bar{J}^P_1 \\[2ex] P^N_j - \rho^{rr} P^N_j + \rho^{nn} \left(\sum_{\zeta<j} P^N_\zeta \right) & , \quad j = \underline{J}^P_1 - 1 = J^0_1 \\[2ex] P^N_j - \rho^{rr} P^N_j - \rho^{nn} P^N_j & , \quad j = 1, \ldots, \underline{J}^0_1 - 1 \end{cases} \tag{4.149}$$

[40] This is the case where the largest possible number of bins are distorted by both types of rigidity in the two periods, i.e. we have the least possible amount of information about the 'notional' probability histogram that satisfy assumptions (4.139)-(4.143).

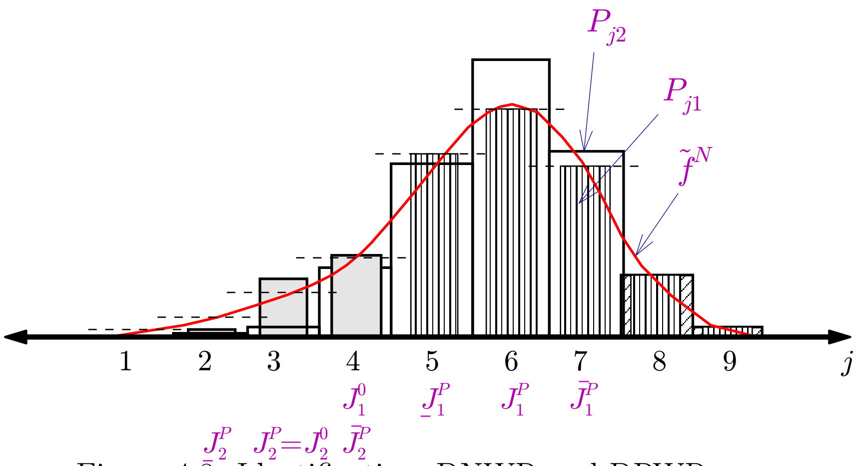


Figure 4.8: Identification, DNWR and DRWR case.

and, in period 2

$$P_{j2} = \begin{cases} P_j^N & , \ j = \bar{J}_2^P + 1 \left(= \underline{J}_1^P\right), \ldots, J \\[2ex] P_j^N + \rho^{rr} \pi_{\bar{J}_2^P - J_2^P} \left(\sum_{\zeta<j} P_\zeta^N\right) & , \ j = \bar{J}_2^P \left(= \underline{J}_1^P - 1 = J_1^0\right) \\[2ex] \left.\begin{array}{l} P_j^N - \\ -P_j^N \rho^{rr} \pi_{\bar{J}_2^P - J_2^P} + \left(\sum_{\zeta<j} P_\zeta^N\right) \rho^{rr} \pi_{j - J_2^P} \\ +\rho^{nn} \left(\sum_{\zeta<j} P_\zeta^N\right) \end{array}\right\} & , \ j = \bar{J}_2^P - 1 = J_2^0 \\[2ex] \left.\begin{array}{l} P_j^N - \\ -P_j^N \rho^{rr} \left(\sum_{\xi=j+1}^{\bar{J}_2^P} \pi_{\xi - J_2^P}\right) + \\ + \left(\sum_{\zeta<j} P_\zeta^N\right) \rho^{rr} \pi_{j - J_2^P} \\ -\rho^{nn} P_j^N \end{array}\right\} & , \ j = \underline{J}_2^P, \ldots, \bar{J}_2^P - 2 \left(< J_2^0\right) \\[2ex] P_j^N - \rho^{nn} P_j^N - \rho^{rr} P_j^N & , \ j = 1, \ldots, \underline{J}_2^P - 1 \end{cases} \tag{4.150}$$

Bins $j = \underline{J}_1^P, \ldots, J$: First we note from the top row of (4.150) that the data from period 2 allows us to estimate directly the bins of the 'notional' probability histogram indexed by $j = \underline{J}_1^P, \ldots, J$ (in Figure 4.8: $j = 5, \ldots, 9$):

$$\left\{P_{j2}\right\}_{j=\bar{J}_2^P+1,\ldots,J} = \left\{P_j^N\right\}_{j=\underline{J}_1^P,\ldots,J} \tag{4.151}$$

Furthermore, from (4.149), we see that the bins of the 'actual' probability histogram in period 1 that contain the support of AID (in Figure 4.8: $j = 5, 6, 7$) are only distorted by DRWR. Therefore combining the information from two periods we can employ the same arguments used in the proof of Theorem 2, to write ρ^{rr} and $\{\pi_q\}_{q \in \mathcal{Q}}$ as functions of identifiable quantities.

Bin $j = \underline{J}_1^P - 1 = J_1^0$**:** With regard to the identification of ρ^{nn}, we will make use of the information about P_{j1} and P_{j2} for $j = J_1^0 \left(= \underline{J}_1^P - 1 = \bar{J}_2^P\right)$ - (in Figure 4.8: $j = 4$), - in addition to the results about ρ^{rr} and $\{\pi_q\}_{q \in \mathcal{Q}}$. Copying from (4.149) and (4.150) we have, for this bin

$$P_{j1} = P_j^N - \rho^{rr} P_j^N + \rho^{nn} \left(\sum_{\zeta < j} P_\zeta^N \right) \tag{4.152}$$

$$P_{j2} = P_j^N + \rho^{rr} \pi_{\bar{J}_2^P - \underline{J}_2^P} \left(\sum_{\zeta < j} P_\zeta^N \right) \tag{4.153}$$

Using

$$\left(\sum_{\zeta < j} P_\zeta^N \right) = 1 - P_j^N - \left(\sum_{\zeta = \underline{J}_1^P}^{J} P_\zeta^N \right) \quad , \quad j = \underline{J}_1^P - 1 \tag{4.154}$$

and (4.151), we can solve (4.153) for P_j^N, where the RHS includes only identifiable quantities:

$$P_j^N = \frac{P_{j2} - \rho^{rr} \pi_{\bar{J}_2^P - \underline{J}_2^P} \left[1 - \left(\sum_{\zeta = \underline{J}_1^P}^{J} P_{\zeta 2} \right) \right]}{1 - \rho^{rr} \pi_{\bar{J}_2^P - \underline{J}_2^P}} \quad , \quad j = \underline{J}_1^P - 1 \tag{4.155}$$

Therefore, the expression we obtain by solving (4.152) for ρ^{nn}

$$\rho^{nn} = \frac{P_{j1} - P_j^N + \rho^{rr} P_j^N}{1 - P_j^N - \left(\sum_{\zeta = \underline{J}_1^P}^{J} P_{\zeta 2} \right)} \quad , \quad j = \underline{J}_1^P - 1 = J_1^0$$

will also include only identifiable quantities on its RHS. ∎

Bins $j = 1, \ldots, \underline{J}_1^P - 2$**:** Given that ρ^{nn} and ρ^{rr} are identifiable, it is straightforward to show that the remaining bins of the 'notional' probability histogram, i.e. with index values $j = 1, \ldots, \underline{J}_1^0 - 2$, are also identifiable (in Figure 4.8: $j = 1, 2, 3$). The easiest way is by solving the expression in the last row of (4.149) for P_j^N

$$P_j^N = \frac{P_{j1}}{1 - \rho^{rr} - \rho^{nn}} \quad , \quad j = 1, \ldots, \underline{J}_1^0 - 2 \tag{4.156}$$

We conclude this section by noting that, in order to check in practice whether the sufficient conditions for identification stated above are satisfied for the available data, one must know the values of $\left\{ J_t^0, \underline{J}_t^P, \bar{J}_t^P \right\}_{t=1,\ldots,T}$. With regard to J_t^0, this is easy to calculate. On the other hand, for $\left\{ \underline{J}_t^P, \bar{J}_t^P \right\}_{t=1,\ldots,T}$ one would have to estimate $E\dot{P}_{ti}^e$ (for all t), which would determine the value of J_t^P, and then make assumptions about the number of bins that, in addition to J_t^P, also contain values of the support of AID.

4.4.3 Heterogeneity

In the model described in Section 4.2, we allowed for heterogeneity in the mean of the notional WGD only across periods. This meant that the mean notional wage growth rate was the same across individuals within a particular period. It also followed that mean heterogeneity coincided with time heterogeneity, therefore the period index t could also be interpreted as a mean-heterogeneity group index.

More generally, and depending on the type of data we are analysing, there may be reasons to believe that mean heterogeneity exists across individuals within periods as well. For example, when dealing with data on individual wage growth rates, then individual characteristics, such as measures of the individual's human capital stock, or firm characteristics, such as the industry the firm is operating in and the extent of unionisation across employees, could be expected to have an influence on the mean wage growth rate that would prevail in the absence of rigidity.[41] On the other hand, when dealing with wage growth data from collective agreements, then both firm and union characteristics could be expected to have an influence on the mean notional wage growth rate.

It is possible to handle this 'generalised' type of mean heterogeneity in the context of the semiparametric approach described earlier by treating the data from each heterogeneity group defined according to a set of characteristics in the same way we treated the data from each period in the earlier analysis.

We recall that, previously, in order to handle mean heterogeneity across periods, we defined T standardised probability histograms, one for each period. Then the contribution of particular observation (y_{ti}) to the likelihood function was determined by the nature of the probability histogram of the period the observation belonged to.

Accordingly, in the presence of more extensive heterogeneity, a standardised probability histogram must be defined for each mean-heterogeneity group. In that case, index t will no longer be the mean-heterogeneity group index. Instead we could define a composite index to map the heterogeneity groups defined by all possible values taken by the vector of covariates $\mathbf{x}_{ti}$ recording mean heterogeneity for individual i from period t, denoted by $\vartheta\left(\mathbf{x}_{ti}\right)$ [42]

$$\vartheta\left(\mathbf{x}_{ti}\right) : \mathcal{X} \equiv \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_\Theta \to \mathbb{N}_+ \tag{4.157}$$

$$\mathbf{x}_{ti} \equiv \left[x_{\theta,ti}\right]_{\theta=1,\ldots,\Theta} \quad , \quad x_{\theta,ti} \in \mathcal{X}_\theta \tag{4.158}$$

Then we could denote the wage growth rate of individual i from period t by $\dot{w}_{\vartheta(\mathbf{x}_{ti}),i}$, and accordingly, the 'discretised' wage growth variable by $y_{\vartheta(\mathbf{x}_{ti}),i}$, and write the likelihood function as follows

$$lnL_2 = \sum_{\vartheta(\mathbf{x}_{ti})} \sum_{i \in \mathcal{I}_{\vartheta(\mathbf{x}_{ti})}} \sum_{j=1}^{J} I_{\left(y_{\vartheta(\mathbf{x}_{ti}),i}=j\right)} ln\left(P_j^N + D_{j,\vartheta(\mathbf{x}_{ti})}^n + D_{j,\vartheta(\mathbf{x}_{ti})}^r\right) \tag{4.159}$$

where $\mathcal{I}_{\vartheta(\mathbf{x}_{ti})}$ is the set of index values of those individuals that belong to heterogeneity group $\vartheta\left(\mathbf{x}_{ti}\right)$.

[41] See for example Altonji and Devereux (2000).

[42] Obviously, if we only have mean-heterogeneity across t, then $\vartheta\left(\mathbf{x}_{ti}\right) : \mathcal{X} \equiv \{1, \ldots, T\} \to \mathbb{N}_+$, and $\vartheta\left(\mathbf{x}_{ti}\right)$ coincides with t.

Two important issues arise with this kind of treatment of heterogeneity in the context of the semiparametric approach just described. One is concerned with the handling of heterogeneity that is continuous in nature. This cannot be explicitly accommodated in the way described above; in that case, $\vartheta(\cdot)$ would be a continuous variable whose range is a subset of $\mathbb{R}$, and thus cannot be mapped into the set of natural numbers. One possible way to handle this would be to group the values of the continuous heterogeneity variable, and then proceed with the analysis using the discrete variable that records these groups of values.

A second issue is concerned with the amount of data that becomes available from each mean-heterogeneity group considered in the analysis as the number of heterogeneity groups increases. Obviously, the number of observations per heterogeneity group diminishes as the number of heterogeneity groups considered increases. This would affect the quality of the estimates of the probability histograms and, consequently, of the rigidity parameters. This takes place in two ways: directly, due to the use of fewer observations to estimate the probability histograms corresponding to each heterogeneity group, as well as indirectly, through its effect on the quality of the estimate of the location parameter used at the standardisation stage. Both these issues introduce limitations to the extent that heterogeneity could be treated by the semiparametric approach, that do not exist for the parametric approach.

We note at this point that other forms of heterogeneity could also be accommodated by the semiparametric approach, for example, with respect to the rigidity parameters (i.e. ρ^n and ρ^r), the wage-setting regime coverage parameters (i.e. p^n and p^r), as well as the shape of the notional distribution.[43] In particular, for the cases of heterogeneity in the rigidity parameters and the parameters that measure the coverage of the wage-setting regimes, this could be accommodated through their parameterisation using functions of the covariates that measure this heterogeneity. For example, ρ^n and ρ^r could be heterogenous with respect to contract duration (in the case of collective agreement data).[44] On the other hand, p^n and p^r might depend on the state of the economy, e.g. p^n could depend on contract duration and anticipated inflation, and p^r could depend on contract duration, anticipated inflation, as well as productivity growth. Following this parameterisation strategy might help to separately identify ρ^n from p^n, and similarly ρ^r from p^r, if different sources of heterogeneity are relevant for each of the parameters in these pairs, and/or different functional forms are used for parameterisation.[45]

[43] Ignoring such types of heterogeneity could lead to biases in the estimated parameters.

[44] In particular, realisation of negative nominal or (anticipated) real wage changes might be more likely for shorter contracts. We note, though, that there may be an issue of endogeneity of contract duration in this context.

[45] This issue warrants further investigation.

4.4.4 Empirical implementation issues

Standardisation

The problem The methodology described in Section 4.4.1 requires the standardisation of the location of the distribution underlying the data by subtracting from each data point a quantity λ_t which is equal to the value of some location parameter of the *notional* distribution, such as the mean or some quantile. However, such population parameter values are typically neither known, nor can they be directly estimated as we do not observe the notional wage growth rates. Therefore, the above methodology is not feasible as it stands.

Solution A way forward is to consider implementing the standardisation using an estimate of such a location parameter - i.e. of the *notional* WGD - that is based on *actual* wage growth rate data. This is feasible as certain location parameters of the actual WGD coincide with the corresponding location parameters of the notional; in particular, this is true for the quantiles of the actual WGD that lie in the region that cannot be distorted in the presence of rigidity. We recall from the discussion in Section 4.2, that the notional and actual WGDs coincide beyond a certain point, which we denote by $\dot{w}_t^0$

$$F_t\left(w\right) = F_t^N\left(w\right) \quad , \quad w = \dot{w}_t^0 \tag{4.160}$$

$$f_t\left(w\right) = f_t^N\left(w\right) \quad , \quad w > \dot{w}_t^0 \tag{4.161}$$

where the critical point $\dot{w}_t^0$ depends on the nature of the rigidity mechanism

$$\dot{w}_t^0 = \begin{cases} 0 & , \quad \text{under DNWR} \\ \max_i\left(\dot{P}_{ti}^e\right) & , \quad \text{under DRWR} \\ \max\left(0, \max_i\left(\dot{P}_{ti}^e\right)\right) & , \quad \text{under DNWR and DRWR} \end{cases} \tag{4.162}$$

Therefore all quantiles of the actual WGD that lie to the right of this point must be identical to the same order quantiles of the notional distribution, i.e.

$$q_{\alpha,t} = q_{\alpha,t}^N \quad , \quad q_{\alpha,t}^N > \dot{w}_t^0 \tag{4.163}$$

where $q_{\alpha,t}$ and $q_{\alpha,t}^N$ are the α'th percentiles of the actual and notional WGDs for period t, respectively. We can therefore proceed to standardise the actual WGDs using as a location parameter the estimate of a quantile, denoted by $\hat{\lambda}_t$, that satisfies

$$\hat{\lambda}_t \in \left\{\hat{q}_{\alpha,t} : q_{\alpha,t} > \dot{w}_t^0 \ \forall t\right\} \tag{4.164}$$

calculated from the available data on actual wage growth rates.[46] A natural estimator is the sample quantile.[47]

Implications It turns out that using $\hat{\lambda}_t$ rather than λ_t means that the methodology described in Section 4.4.1 is valid only asymptotically.

To see why this is the case, we first note that when standardisation takes place using $\hat{\lambda}_t$, then the relevant location-standardised wage growth rate variable is not $\tilde{w}_{ti}$, defined in equation (4.31), but instead w^*_{ti}, given by

$$w^*_{ti} \equiv \dot{w}_{ti} - \hat{\lambda}_t \tag{4.165}$$

Also $\hat{\lambda}_t$, being an estimator of the population quantity λ_t, relates to the latter in the following way

$$\hat{\lambda}_t = \lambda_t + e_t \tag{4.166}$$

$$e_t \sim q(\cdot) \tag{4.167}$$

where e_t is the estimation error whose properties depend on the properties of the sample quantile as an estimator of population quantiles. It follows that the relationship between w^*_{ti} and $\tilde{w}_{ti}$ is given by

$$w^*_{ti} = \tilde{w}_{ti} - e_t \tag{4.168}$$

Suppose now that we have a particular realisation of data that provides us with a particular estimate of λ_t such that the error of estimation is equal to a particular value, denoted by e. Assuming $\lambda_t > 0$ then, from (4.165), if the estimation error is positive, therefore $\hat{\lambda}_t > \lambda_t$, then using $\hat{\lambda}_t$ rather than λ_t has a result the shift of the actual WGD to the left by a distance e in excess of the correct amount. Accordingly, the opposite happens when the estimation error is negative and $\lambda_t > 0$. In both cases the standardisation of the *location* is not correctly implemented, however the presence of a non-zero estimation error has no effect on the shape of incorrectly standardised distribution.[48]

We then define the probability of observing w^*_{ti} in bin j conditional on $e_t = e$ as follows

$$P^*_{jt}(e) \equiv \Pr\left(\eta_{j-1} \leq w^*_{ti} < \eta_j | e_t = e\right) \tag{4.169}$$

[46]In practice $\max_i(\dot{P}^e_{ti})$ is not easy to be determined as $\dot{P}^e_{ti}$ is unobservable. However the estimate of $E\dot{P}^e_{ti}$, which can be obtained (see Chapter 2), can provide a lower bound. In order to check the robustness of the results one can, in practice, consider producing estimates of the rigidity parameters using several choices of quantiles whose estimates are used as standardisation parameters.

[47]We note at this point that (an estimate of) the mean of the actual WGD is not a suitable candidate for a location parameter for the standardisation, since, typically, it is higher than the mean of the notional WGD due to the shift of mass to the right caused by the presence of rigidity.

[48]We can see this by noting that the relationship between the correctly standardised wage growth variable $\tilde{w}_{ti}$ and the incorrectly standardised wage growth variable w^*_{ti} is given by $w^*_{ti} = \tilde{w}_{ti} - e$. From this it follows that relationship between the PDF of the distribution of w^*_{ti} conditional on $e_t = e$, denoted by $f^*_t(w^*_{ti}|e_t = e)$ with the (unconditional) distribution of $\tilde{w}_{ti}$ is given by $f^*_t(w|e_t = e) = \tilde{f}_t(w + e)$ where the PDF on the right hand side of this expression is defined with respect to w, and e is a constant.

Then, using (4.168), we can write

$$P^*_{jt}(e) = \Pr(\eta_{j-1} \leq \tilde{w}_{ti} - e_t < \eta_j | e_t = e) = \Pr(\eta_{j-1} \leq \tilde{w}_{ti} - e < \eta_j) = \\ = \Pr(\eta_{j-1} + e \leq \tilde{w}_{ti} < \eta_j + e) \equiv P_{jt}(e) \tag{4.170}$$

that is, we can view the standardisation that is based on an estimate of the location parameter as one that is based on the true value of that parameter, which is, however, followed by the stage of discretisation where the location of the bins is not the same across periods. In particular, they are misplaced by a distance e in the opposite direction of the error committed under the standardisation.

Using (4.170) we can thus write the unconditional probability of observing w^*_{ti} in bin j as a mixture of probabilities defined with respect to the correctly standardised distribution

$$\begin{aligned} P^*_{jt} &= \Pr(\eta_{j-1} \leq w^*_{ti} < \eta_j) \\ &= \int_e \Pr(\eta_{j-1} \leq w^*_{ti} < \eta_j | e_t = e)\, q(e)\, de \\ &= \int_e P_{jt}(e)\, q(e)\, de \end{aligned} \tag{4.171}$$

where the mixing distribution is the distribution of the estimation error. If $\hat{\lambda}_t$ consistently estimates λ_t, then by definition the distribution of the estimation error will become increasingly concentrated around the value zero, and P^*_{jt} will converge in probability to $P_{jt}(0)$, which is equal to P_{jt}. Therefore, although the log-likelihood function in (4.76) is misspecified in finite samples, it is still valid asymptotically. It improves, as an approximation to the correct log-likelihood function, the better the estimate of the location parameter we use; this, of course, depends on the amount of data that is available from each period.

In Section 4.5 we use simulations to examine the impact that the use of $\hat{\lambda}_t$ rather than λ^N_t has on the estimates of ρ^n and ρ^r. In that respect, it is particularly interesting to investigate the implications from using the sample median ($\hat{m}_t$) as a standardisation parameter under the presence of DRWR, which, in a way, is the 'default' choice given its widespread use in studies that investigate the presence of DNWR,[49] given that it is not unlikely to have $m_t < \max(\dot{P}^e_{ti})$ for some t.

Discretisation

The definition of bins for the implementation of the standardisation stage could be seen as involving three distinct decisions; one about the choice of the width of the bins, i.e. $\eta = \eta_j - \eta_{j-1}$, another about the exact location of the bins, i.e. the exact values taken by the bin endpoints, and one about the number of bins.[50]

With respect to the choice of bin width, the basic concern is about avoiding undersmoothing or oversmoothing of the probability histogram, which would result from choos-

[49]E.g. Kahn (1997) and Beissinger and Knoppik (2001).

[50]Given the choice of bin width, the decision about their location simplifies to the choice of location of any of the bin endpoints, such as η_0.

ing a 'too small' or a 'too large' η, respectively. In either case, the shape of the histogram will not accurately represent the main features of the shape of the actual wage growth distribution, including the size of the rigidity distortions, if any. According to the literature on non-parametric density estimation,[51] this choice could be made to satisfy some optimality criterion that would, in general, depend on the sample size. This implies that the optimal bin width will be the same across periods, as is required for the implementation of the semiparametric approach, only if the respective sample sizes are the same. However, this is not, typically, the case, therefore no single bin width can be chosen optimally for all samples. A solution to this could be to choose the optimal bin width according to the average sample size.[52] With respect to the second choice (i.e. location), the statistics literature does not provide any optimality rules, although, this decision could have an effect on the shape of the probability histogram.[53] [54]has been to locate the bins in such a way that point $\tilde{w}_{ti} = 0$ is located in the middle of the bin that contains it, i.e.

$$\eta_{j-1} = -\eta_j \quad , \quad j : \{0\} \in B_j$$

Given the definitions of bins, the number of bins that compose each histogram will depend on the spread of the data within samples. In practice we could choose the number of bins so that they cover satisfactorily the range of values in the sample with the biggest spread. The choice of index values, i.e. the values taken by j, can be made arbitrarily as it has no practical significance, except for notational convenience.

Location and spread of the anticipated inflation distribution

In order to implement the semiparametric method we would have to specify the value of J_t^P for all t, i.e. the index value for the bin that contains the standardised value of mean anticipated inflation rate in each t. We recall that, depending on the standardisation method used, the standardised value of mean anticipated inflation rate is given by

$$E\tilde{P}_{ti}^e = E\dot{P}_{ti}^e - \lambda_t \tag{4.172}$$

or, more likely in practice, by

$$EP_{ti}^{*e} = E\dot{P}_{ti}^e - \hat{\lambda}_t \tag{4.173}$$

This calculation requires knowledge of $E\dot{P}_{ti}^e$ which, typically, is unknown since $\dot{P}_{ti}^e$ is not observable. In Chapters 2 and 3 we dealt with this issue using inflation forecasts in place of $E\dot{P}_{ti}^e$, obtained from an estimated time series model of inflation.

We note at this point that there is no guarantee that the identifying condition given

[51]See, for example, Wasserman (2006).

[52]In Section 4.5 we use simulations to investigate the implications of the sample size, given bin width, for the properties of the rigidity parameter estimates.

[53]In fact, this effect underlies one of the problems with using probability histograms as approximations to the density function, as the shape is depended on the choice of histogram location. See, for example, Silverman (1986).

[54]See for example Kahn (1997), Beissinger and Knoppik (2001) and Holden and Wulfsberg (2007).

in (4.69), regarding the location of $E\tilde{P}^e_{ti}$ or EP^{*e}_{ti} in the bin that contains it, will always be satisfied.[55] It is clear, though, that by defining the bin width as small *as possible*, we minimise the extent of this problem.

In addition to J_t^P, the model also requires the specification of any additional bins that include values of anticipated inflation rates. In terms of the notation introduced earlier, this refers to the specification of the set $\mathcal{Q}$, given the value of J_t^P. In practice this could be determined by the data, starting off with a certain specification of $\mathcal{Q}$, and then testing for the statistical significance of the π_q's to decide which terms to keep.[56]

Imposing restrictions on the parameters in the log-likelihood function

All the unknown parameters to be estimated through maximisation of the log-likelihood function are probabilities, and therefore take values in the interval $[0, 1]$. Furthermore, the respective sums of the notional and AID probabilities must be equal to one. By default, the maximisation of the log-likelihood function takes place over the real line for each of the unknown parameters, and, as the definition of the log-likelihood stands in (4.76), there is no guarantee that the estimates obtained will satisfy these restrictions. In order to make sure that this will be the case, we opt for the re-parameterisation of the likelihood function by replacing each of the unknown parameter with the logistic function of an auxiliary unknown parameter, i.e.

$$\rho \equiv \frac{\exp(\rho_*)}{1 + \exp(\rho_*)} \tag{4.174}$$

$$p_j = \frac{\exp(p_{j*})}{\sum_j \exp(p_{j*})} \tag{4.175}$$

where ρ is used generically to denote ρ^{nn} and ρ^{rr}, and ρ_* the corresponding auxiliary parameters, and, similarly, p_j is used generically to denote P_{jt}^N and π_q, and p_{j*} the corresponding auxiliary parameters. We note that for the latter parameterisation we need to normalise one of the auxiliary parameters to be equal to zero.[57]

4.4.5 Link to the location-histogram approach

It is possible to motivate the location-histogram approach described in Chapters 2 and 3 from within the framework developed so far.

Using (4.52), we define the vector of *proportions* or *relative frequencies* of standardised

[55]Further study of the implications of such failure is required.

[56]If our estimate of $E\dot{P}^e_{ti}$ were relatively accurate, then we would expect to find π_0 to be statistically significant. Therefore the issue is mainly about the remaining values of q.

[57]Given that the logistic function has as domain the real line and range the $[0, 1]$ interval, the maximisation of the log-likelihood function w.r.t. to the auxiliary parameters can take place over the $(-\infty, +\infty)$ interval, and, at the same time, the estimates of the original parameters will be proper probability estimates in the $[0, 1]$ interval. The choice of the logistic function was made for computational convenience; any other function with the same domain and range could have been used instead.

observations, in each period t, that fall in each bin, as follows

$$\hat{\mathbf{P}}_t \equiv \left[\hat{P}_{jt}\right]_{j=1,\ldots,J} = \left[\frac{1}{n_t}\sum_{i=1}^{n_t} d_{tij}\right]_{j=1,\ldots,J} = \frac{1}{n_t}\sum_{i=1}^{n_t} \mathbf{d}_{ti} = \frac{\mathbf{d}_t}{n_t} \tag{4.176}$$

It, then, follows from (4.52) that

$$\mathbf{d}_t = \sum_{i=1}^{n_t} \mathbf{d}_{ti} \sim Multinomial\left(n_t, \mathbf{P}_t\right) \tag{4.177}$$

and, therefore,

$$n_t \hat{\mathbf{P}}_t = \mathbf{d}_t \sim Multinomial\left(n_t, \mathbf{P}_t\right) \tag{4.178}$$

Given data on wage growth rates from a particular period t, we can evaluate the vector of proportions, and the resulting value, denoted by $\hat{\mathbf{p}}_t = \left[\hat{p}_{jt}\right]_{j=1,\ldots,J}$, can be seen as a realisation of $\hat{\mathbf{P}}_t$. Expression (4.178) can then be used to construct the likelihood function for the sample of vectors of proportions $\{\hat{\mathbf{p}}_t\}_{t=1,\ldots,T}$ as follows

$$L_3 = \prod_{t=1}^{T}\left[\frac{n_t!}{\prod_{j=1}^{J}\left(n_t\hat{p}_{jt}\right)!}\prod_{j=1}^{J} P_{jt}^{(n_t\hat{p}_{jt})}\right] \propto L_2 \tag{4.179}$$

where P_{jt} is given by expressions (4.73)-(4.75). Alternatively, we could follow a regression approach, discussed by Zellner and Lee (1965) that involves the estimation of a system of J equations

$$\hat{\mathbf{P}}_t = E\hat{\mathbf{P}}_t + \epsilon_t = \mathbf{P}_t + \epsilon_t \tag{4.180}$$

with T observations on each equation, where the dependent variable for equation j is $\hat{P}_{jt}$. As above, P_{jt} is given by expressions (4.73)-(4.75).[58] This estimation approach is identical to the one followed in Chapters 2 and 3. As we saw there, the variance-covariance matrix of the vector of dependent variables is not spherical, and a FGLS regression approach provides asymptotically efficient estimates of the unknown parameters.[59]

This regression approach could be described as an application of the Minimum Chi-Squared Estimation (MCSE) method for multivariate grouped discrete data. It is an extension of the work by Berkson (1955), who proposed this regression approach for the univariate case.[60] Zellner and Lee (1965) have shown that the minimum chi-squared estimators of the parameters in a linear regression system like the one in (4.180) are consistent

[58] The regression equations in this system will be non-linear, and therefore the implementation of the non-linear least squares estimation procedure is called for. We note that in their work Zellner and Lee (1965) dealt with the linear probability model, as was the case with the models considered in Chapters 2 and 3. However, their results extend to the case studied here as well.

[59] The estimation approach described in Chapter 2 follows closely, but not exactly, the original approach of Kahn (1997). In her case, a SUR-type estimator of the model in (4.180) was implemented, while in Chapter 2 we derive the analytic form of the relevant variance-covariance matrix, whose elements are shown to be functions of the probabilities P_{jt}, and the yearly sample sizes n_t. The P_{jt}'s are then consistently estimated in a first stage – and so is the variance-covariance matrix – using the corresponding sample proportions, and in a second stage the FGLS estimator is implemented.

[60] Maddala (1983) provides a review of the literature on the application of the MCSE principle for the estimation of discrete endogenous models.

and asymptotically Normal. More generally, it has been shown that the MCSE principle applied to discrete endogenous variable models, including the logit and probit models, can produce asymptotically efficient estimators as well, and is, thus, equivalent to the ML approach (see Amemiya (1976)).

On the other hand, in finite samples the MCSE principle produces good estimates as long as two conditions are met (see Domencich and McFadden (1975)); (a) a 'sufficient' number of observations from each heterogeneity group in the population, according to which 'population cells' are defined, are available, and, (b) the continuous (explanatory) variables measuring population heterogeneity are 'appropriately discretised' in order to define the necessary population cells for the implementation of the MCSE approach. According to Domencich and McFadden (1975)), in principle, the cells based on the discretised continuous explanatory variables should be redefined as the sample size increases so that the number of observations in each cell increases and, at the same time, the range of variation of the values taken by the relevant continuous explanatory variable for the observations that fall in each cell decreases.[61] For the proportional DWR model considered here, population heterogeneity only exists across years. This means that, on the one hand, there is no ambiguity with respect to the satisfaction of condition (b) above as this is discrete in nature. On the other hand, one would expect that, in practice, the number of observations from each year n_t - i.e. the cell size, for this case, - is large enough so that condition (a) above is also met.[62] Therefore for the particular model examined here, the MLE and MCSE approaches are likely to yield very similar results in typical finite sample conditions.[63]

4.5 Simulations

In this section we present the results from several simulation exercises which aim to investigate some aspects of the finite sample properties of the estimators suggested by the methodology just presented. The parameters which are allowed to vary across experiments

[61] The amount of observations from each cell that is deemed to be 'sufficient' is not fixed. Domencich and McFadden (1975) present simulation results for a simple univariate logit model estimated by ML and MCSE, with one continuous explanatory variable whose values were grouped in several different ways, and where the sample size varied between 30 to 240 observations. According to their results, the optimal cell size for the experiments they considered ranged between 3 to 8 observations, the bigger value corresponding to the larger sample sizes. Further they show that the way the discretisation of the continuous explanatory variable takes place is crucial, as it can lead to large biases in the estimates obtained.

[62] For example, for the contract data used in Chapters 2 and 3 we had approximately 500 observations per year, on average.

[63] We note that if we were to consider extensions to this model to account for within-year heterogeneity, as discussed in Section 4.4.3, conditions (a) and (b) would progressively become more difficult to satisfy, and the appeal of the MCSE would soon diminish. For example, in the case of discrete independent variables measuring heterogeneity, the number of heterogeneity groups is given by all possible configurations of their values, which tends to increase with the power of the number of variables (e.g. if K variables, then 2^K heterogeneity groups/cells.) More generally speaking, the MCSE approach seems to better suit experimental data, where experiment conditions are likely to be measured by discrete variables, and where repetitions of the experiments is feasible. On the other hand, this is not typically the case with survey data, which typically originate from populations characterised by continuous-type heterogeneity, and where the total sample sizes cannot much the requirements of condition (a).

are the number of periods samples are available from (T)[64], the size of these samples (n_t), the degree of heterogeneity across samples, the intensity of rigidity (i.e. the size of ρ^n and/or ρ^r), the spread of the AID - when we allow for DRWR, - and the standardisation method employed for the implementation of the semiparametric approach. The results reported include the sample mean and standard deviation of the estimates of the rigidity parameter(s), as well as the Root Mean Squared Error (RMSE) of estimation. For comparison purposes, we also report the corresponding results obtained from the application of the parametric approach on simulated data from the same data generating process (DGP).

First we consider the case where the data is only contaminated by DNWR, then the case where the data is only contaminated by DRWR, and conclude with the case where both types of rigidity are present.

4.5.1 Data contaminated by DNWR

Basic setup

The DGP In terms of the model described in Section 4.2, the model for the DGP in this case emerges from setting $p^n = 1$. The relationship between the generated 'actual' $(\dot{w}_{ti})$ and the corresponding 'notional' $(\dot{w}_{ti}^N)$ wage growth rates then simplifies to the following

$$\dot{w}_{ti} = \begin{cases} \dot{w}_{ti}^N & , \quad \dot{w}_{ti}^N \geq 0 \\ (1-\delta_{ti}^n)\,\dot{w}_{ti}^N & , \quad \dot{w}_{ti}^N < 0 \end{cases} \tag{4.181}$$

where

$$\delta_{ti}^n = \begin{cases} 1 \quad , \quad \delta_{ti}^{*n} \leq \rho^n \\ 0 \quad , \quad \delta_{ti}^{*n} > \rho^n \end{cases} \quad , \quad \delta_{ti}^{*n} \overset{IID}{\sim} U(0,1) \tag{4.182}$$

To simulate the data, the notional wage growth rates are sampled from the Normal distribution

$$\dot{w}_{ti}^N = \mu_{Nt} + \epsilon_{ti}^N \tag{4.183}$$

$$\epsilon_{ti}^N \overset{IID}{\sim} N\left(0, \sigma_N^2\right) \quad , \quad \sigma_N \simeq 4.1 \tag{4.184}$$

where the variance is the same across samples, but the mean is allowed to be different.[65] Specifically we allow this to behave as a Uniformly distributed random variable, with 'low' mean $(E\mu_{Nt} = E\mu_L = 2)$ for the first τ samples , and 'high' mean $(E\mu_{Nt} = E\mu_H = 14)$ for the remaining $T-\tau$ samples

$$\mu_{Nt} = \begin{cases} \mu_L \overset{IID}{\sim} U(0,4) & , \quad 1 \leq t \leq \tau \\ \mu_H \overset{IID}{\sim} U(12,16) & , \quad \tau < t \leq T \end{cases} \tag{4.185}$$

With this configuration we can control the number of samples which come from undistorted

[64] Or, more generally, the number of mean-heterogenous groups in the population.

[65] Therefore, the distributions underlying the generated samples, indexed by t, have the same shape but differ in their location, which is determined by μ_{Nt}.

'actual' distributions[66] - those that correspond to the high-mean 'notional' distributions, - and from rigidity-contaminated 'actual' distributions - those that correspond to the low-mean 'notional' distributions.[67] The values of ρ^n and τ that are required to complete the specification of the DGP are left as parameters for the simulation designs, and are discussed later in the text.

Details of implementation of semiparametric approach For the implementation of the semiparametric approach we need to specify the value of the location parameter that is used for the standardisation, and the details regarding the partition of the support of the notional distribution for the implementation of the 'discretisation' stage.

Regarding the second choice, for all simulation exercises executed we have set the bin width ($= \eta \equiv \eta_j - \eta_{j-1}$) equal to 1%, the number of bins ($= J$) equal to 17, and $\eta_0 = -8.5$. Therefore, the bin limits take values from the set

$$\eta_j \in \{-8.5, -7.5, \ldots, -0.5, 0.5, \ldots, 7.5, 8.5\} \tag{4.186}$$

Given our choice of value for $\sigma_N \simeq 4.1$, the probability histograms resulting from this specification of discretisation cover above 96% of the probability mass of the underlying continuous distribution. With regard to the choice of standardisation parameter, this is left as a simulation parameter; it could be either equal to the true value of the notional median, or to the sample median of the 'actual' wage growth data.

For the construction of the log-likelihood function we note that, in the presence of only DNWR, the relationship between the 'actual' (rigidity-contaminated) and 'notional' probability histograms simplifies to the following:[68]

$$P_{jt} = P_{jt}^n = \begin{cases} P_j^N & , \quad j > J_t^0 \\ P_j^N - \rho_j^n P_j^N & , \quad j < J_t^0 \\ P_j^N + \rho^n \sum_{\zeta<j} P_\zeta^N & , \quad j = J_t^0 \end{cases} \tag{4.187}$$

Therefore the likelihood function for the sample of discrete random variables[69]

$$\{y_{ti}\}_{t=1,\ldots,T,\; i=1,\ldots,n_t} \tag{4.188}$$

is the following:

$$L_2 = \prod_{t=1}^{T}\prod_{i=1}^{n_t}\prod_{j=1}^{J}\left[P_j^N - \left(\rho_j^n P_j^N\right) I_{\left(y_{ti}<J_t^0\right)} + \left(\rho^n \sum_{\zeta<j} P_\zeta^N\right) I_{\left(y_{ti}=J_t^0\right)}\right]^{I_{(y_{ti}=j)}} \tag{4.189}$$

[66] That therefore coincide with the corresponding 'notional' WGDs.

[67] The high-mean distributions have virtually zero probability mass below the point zero (at most, this quantity is less than 0.2%, which is when $\mu_{Nt} = 12$). On the other hand, the low-mean distributions always have a substantial amount of mass below the point zero, which varies between 16.4% (when $\mu_{Nt} = 4$) and 50% (when $\mu_{Nt} = 0$).

[68] A simplification of the expression in (4.66).

[69] These are defined according to (4.45). Since the $\dot{w}_{ti}^N$'s are drawn independently across t and i, the y_{ti}'s will also be independent across t and i.

Details of implementation of parametric approach Adapting expression (4.16) to reflect the choice of Normality, the PDF of the 'actual' (rigidity-contaminated) wage growth rates will be given by

$$f_t\left(\dot{w}\right) = \begin{cases} \frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right) & , \quad \dot{w} > 0 \\ \left(1-\rho^n\right)\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right) & , \quad \dot{w} < 0 \\ \rho^n \Phi\left(-\frac{\mu_{Nt}}{\sigma_N}\right) & , \quad \dot{w} = 0 \end{cases} \tag{4.190}$$

It follows that the likelihood function for the sample of simulated 'actual' wage growth rates

$$\left\{\dot{w}_{ti}\right\}_{t=1,\ldots,T,\; i=1,\ldots,n_t} \tag{4.191}$$

which are drawn independently from each other, will be given by[70]

$$\begin{aligned} L_1 = \prod_{t=1}^{T}\left[\prod_{\dot{w}_{it}<0}\left(1-\rho^n\right)\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right)\right]\times\left[\prod_{\dot{w}_{it}=0}\rho^n\Phi\left(-\frac{\mu_{Nt}}{\sigma_N}\right)\right]\times \\ \times\left[\prod_{\dot{w}_{it}>0}\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}-\mu_{Nt}}{\sigma_N}\right)\right] \end{aligned} \tag{4.192}$$

Experiment design

The parameters left to vary for the design of the simulation exercises are ρ^n and τ, that relate to the specification of the DGP, the value of the location parameter λ,[71] which is relevant for the implementation of standardisation in the application of the semiparametric approach, and the values of T and n_t, which determine the amount of data.

In Table 4.2 we have collected the values taken by these parameters for the four sets of simulation exercises that are executed. First, a benchmark case, referred to as the 'base case', where we consider 'middle' values for all the parameters; ρ^n is set equal to 0.5, an equal number of samples from each heterogeneity group is specified according to the rule

$$\tau = \begin{cases} T/2 & , \quad \text{T is even} \\ \left(T-1\right)/2 & , \quad \text{T is odd} \end{cases} \tag{4.193}$$

and the standardisation location parameter is set equal to the true value of the median of the notional WGD.[72] The number of samples T takes values from the set $\{5, 10, 20, 40, 80\}$, and for each of these values, the sample size n_t takes values from $\{63, 125, 250, 500\}$.

Then we consider three sets of variations relative to the base case: In Variation 1 we deviate from the base case and allow ρ^n to take a 'low' value ($= 0.2$) for a subset

[70] It is interesting to note from the use of the Standard Normal PDF and CDF in this expression, that the implementation of the parametric approach also involves the 'standardisation' of the data with respect to both the location and scale of the notional WGD. In contrast to the semiparametric approach though, the location and scale parameters are estimated jointly with the rigidity parameters rather than separately. The joint identification of the location and scale parameters of the notional distribution and the rigidity parameter is, of course, possible due to the stronger parametric assumptions that have been made.

[71] We will use λ generically to indicate both λ_t (population value) and $\hat{\lambda}_t$ (estimate).

[72] Given the symmetry of the Normal distribution this is equal to the notional mean.

of experiments and a 'high' value (= 0.8) for another, as opposed to the 'middle' value (= 0.5) used in the base case. In Variation 2 we specify, for a subset of experiments, the minimal degree of heterogeneity across samples that would allow the identification of the rigidity parameter ($\tau = 1$, i.e. only one sample from a DNWR-distorted distribution,) and, for another subset, we allow all samples to be distorted by DNWR ($\tau = T$). Finally, in Variation 3 we set the location parameter equal to the sample median; in this case the DGP is the same as in the base case, but we consider an alternative configuration for the semiparametric approach. For each variation, T and n_t take values as in the 'base' case.

All experiments are replicated 100 times. Given this small number of replications, the simulation results should be interpreted only as indicative.[73]

Results

The results from the execution of the simulation experiments appear in Tables 4.2-4.7. The layout of each table is as follows:

Each row reports the results from a particular experiment; these differ from each other only with respect to the values of T and n_t, which are reported in columns 1 and 2.[74] Column 3 reports the total number of observations that results from pooling the samples across periods (= $T \times n_t$). The results that refer to the estimates obtained from the implementation of the semiparametric approach are reported in columns 4-7, and include the mean and standard deviation of the estimates of ρ^n across replications, the Root Mean Squared Error of estimation, and the average - across replications - of the total number of observations that are used to obtain the semiparametric estimates (= $\overline{T \times n_t}$).[75] The corresponding values for the mean, standard deviation and Root Mean Squared Error of estimation from the implementation of the parametric approach are reported in columns 8-10.[76]

The base case The results from the implementation of the semiparametric approach for the 'base' case (Table 4.2) suggest that the mean of the distribution of the estimator is close to the true value, even when estimation is based on a relatively small number of observations. At the same time the estimator appears to be positively biased as, for all combinations of T and n_t, the mean is larger than the true value.

[73]Stronger caution is warranted in the cases examined in the subsequent two sections where the number of replications is only 50.

[74]For ease of reference the values of the remaining simulation parameters appear at the bottom of each table.

[75]This number if smaller than the total number of observations reported in column 3, as for the implementation of the semiparametric approach we drop all observations that lie outside the bins defined at the discretisation stage. Also, the number of observations is different across replications due to the sampling variability, that is why we report its average. On the other hand, the number reported in column 3 is the number of observations used for the parametric estimation.

[76]We do not report the parametric results in Table 4.7, where we report the results obtained from the implementation of the semiparametric approach when an alternative standardisation method to the one used in the base case is considered. This is because the DGP for this case is the same as in the base case and, therefore, the relevant parametric results are those reported in Table 4.2.

At the same time the estimator appears to be consistent, as the value of the RMSE decreases when the total number of observations used for the estimation increases; this is true both when T is kept fixed and n_t increases, and also when the reverse takes place.

On the other hand, keeping the total number of observations fixed, then increasing T, while at the same time decreasing n_t, does not appear to have a clear effect on the RMSE. This suggests that there in no penalty incurred for increasing the number of heterogeneity groups in the model estimated (presumably, as long as the sample size is preserved above a certain value).

Finally, as one would expect, the results for the parametric approach appear to be better than those for the semiparametric approach, when judged in terms of the values of the RMSE. This, also, appears to converge faster to zero than the corresponding statistic from the semiparametric case. Furthermore, the parametric estimator appears to be unbiased.

Variation 1: vary ρ^r Going from $\rho^n = 0.5$ to $\rho^n = 0.2$ (see Table 4.3) does not seem to alter the overall behaviour of the semiparametric estimator, even though for this case the size of the distortions in the height of the probability histogram bins is less than half of the corresponding size in the 'base' case.

On the other hand, the RMSE in the case where we set $\rho^n = 0.8$ (see Table 4.4) appears to be smaller across all combination of T and n_t relative to the base case; this appears to be due to smaller variability in the estimates in the former case, something that is reflected in the smaller values taken by its standard deviation.

Variation 2: vary extent of mean heterogeneity The variation in the number of samples that come from distorted distributions ($= \tau$) appears to affect the variance of the estimator, which appears to decrease the more samples are drawn from distorted distributions (see Tables 4.5 and 4.6). At the same time, there is no clear evidence that this affects the mean. Therefore the RMSE appears to be smaller in the case where $\tau = T$ relative to the base case, and the converse is true for the case where $\tau = 1$. A possible explanation is that the increase in the number of 'distorted' samples increases the amount of information about the size of the distortions in the pooled sample, and therefore the rigidity parameter can be estimated more accurately.

Another interesting result is that, in the case where $\tau = 1$, the parametric estimates seem to diverge from the true value at $T = 80$, and as we increase the sample size.

Variation 3: vary standardisation parameter The results obtained from changing the standardisation method to one using the sample median of the 'actual' data rather than using its true value do not appear to change significantly; except, maybe, that the mean of the estimates in the former case appears to be closer to the true value (see Table 4.7). According to the discussion in Section 4.4.4, this were to be expected since the design of the simulations ensures that $\hat{m}_t > 0$ for all t, therefore condition (4.164) is satisfied.

4.5.2 Data contaminated by DRWR

Basic setup

The DGP In terms of the model described in Section 4.2, the model for the DGP in this case emerges from setting $p^r = 1$. The relationship between the generated 'actual' ($\dot{w}_{ti}$) and the corresponding 'notional' ($\dot{w}_{ti}^N$) wage growth rates then simplifies to the following

$$\dot{w}_{ti} = \dot{w}_{ti}^r = \begin{cases} \dot{w}_{ti}^N & , \quad \dot{w}_{ti}^N \geq \dot{P}_{ti}^e \\ \delta_{ti}^r \dot{P}_{ti}^e + (1 - \delta_{ti}^r)\, \dot{w}_{ti}^N & , \quad \dot{w}_{ti}^N < \dot{P}_{ti}^e \end{cases} \tag{4.194}$$

where

$$\delta_{ti}^r = \begin{cases} 1 & , \quad \delta_{ti}^{*^r} \leq \rho^r \\ 0 & , \quad \delta_{ti}^{*^r} > \rho^r \end{cases} \quad , \quad \delta_{ti}^{*^r} \overset{IID}{\sim} U(0,1) \tag{4.195}$$

To simulate the data, the notional wage growth rates are sampled from the Normal distribution

$$\dot{w}_{ti}^N = \mu_{Nt} + \epsilon_{ti}^N \tag{4.196}$$

$$\epsilon_{ti}^N \overset{IID}{\sim} N\left(0, \sigma_N^2\right) \quad , \quad \sigma_N \simeq 4.1 \tag{4.197}$$

where the variance is the same across samples, but the mean is allowed to vary;[77] specifically, it can take one of two values, depending on the value of parameter τ

$$\mu_{Nt} = \begin{cases} \mu_H = 18 & , \quad 1 \leq t \leq \tau \\ \mu_L = 10 & , \quad \tau < t \leq T \end{cases} \tag{4.198}$$

In addition to simulating the notional wage growth rates, for this case we also need to simulate the values of the anticipated inflation rates, which vary across t and i. These are also sampled from the Normal distribution

$$\dot{P}_{ti}^e = \mu_{\dot{P}t} + \epsilon_{ti}^P \tag{4.199}$$

$$\epsilon_{ti}^P \overset{IID}{\sim} N\left(0, \sigma_{\dot{P}}^2\right) \tag{4.200}$$

where both the mean and standard deviation are fixed across t. The mean, in particular, takes the value

$$\mu_{\dot{P}t} = 14 \quad , \quad \text{all } t \tag{4.201}$$

across all simulations, while the standard deviation $\sigma_{\dot{P}}$ is treated as a simulation parameter and discussed further later. With these assumptions, the anticipated inflation rates are IID across samples; particularly, the location of the AID will be fixed across t. Furthermore, given the assumptions about the DGP of the notional wage growth rate, and in particular, about the mean of the notional WGD, the centre of the AID will lie to the left of the centre/median of the notional WGD for the first τ samples, and to its right for the remaining

[77] As in the case of DNWR, discussed in the previous sub-section, the distributions underlying the generated samples of notional wage growth rates, indexed by t, have the same shape but differ in their location, which is determined by μ_{Nt}.

$T - \tau$ samples. Consequently, only a part of the left tail of those 'actual' WGDs that correspond to the first τ samples will be distorted by DRWR, while a much bigger part - extending beyond the median of the underlying 'notional' WGD - of the 'actual' WGDs that correspond to the remaining $T - \tau$ samples will be distorted.[78]

To complete the specification of the DGP one would have to specify the values of ρ^r and τ; the value of τ is set fixed across all simulations, and is given by the rule in (4.193), while ρ^r is treated as a simulation parameter.

Details of implementation of semiparametric approach With regard to the implementation of the semiparametric approach, the location parameter used for the standardisation is treated as a simulation parameter. Here, for this purpose, apart from using the true value of the median of the notional WGD, we also consider estimates of higher order quantiles. On the other hand, the details regarding the partition of the support of the notional distribution - for the implementation of the 'discretisation' stage - are the same as in the case of DNWR discussed in the previous sub-section. The only modification concerns the value of η_0, which is chosen to be bigger than -8.5 when the standardisation parameter is an estimate of a higher order quantile than the median. This adjustment is necessary, otherwise if we were to define the bins according to (4.186) for those cases as well, then we would omit a significant amount of data points originating from the left tail of the 'actual' WGD.

With respect to the construction of the log-likelihood function in this case, we note that, in the presence of only DRWR, the relationship between the 'actual' (rigidity-contaminated) and 'notional' probability histograms simplifies to the following:[79]

$$P_{jt} = P^r_{jt} = \begin{cases} P_j^N & , \quad j > \bar{J}_t^P \\ P_j^N - \rho_j^r P_j^N & , \quad j < \underline{J}_t^P \\ \begin{array}{l} P_j^N - P_j^N \left(\sum_{\xi > j} \pi_{\xi - \underline{J}_t^P} \right) \rho^r + \\ + \left(\sum_{\zeta < j} P_\zeta^N \right) \pi_{j - \underline{J}_t^P} \rho^r \end{array} & , \quad \underline{J}_t^P \leq j \leq \bar{J}_t^P \end{cases} \tag{4.202}$$

where the values of $\underline{J}_t^P$ and $\bar{J}_t^P$ depend on the choice of value for $\sigma_{\dot{P}}$. It follows that the likelihood function for the sample of the discrete random variables

$$\{y_{ti}\}_{t=1,\ldots,T,\; i=1,\ldots,n_t} \tag{4.203}$$

defined according to (4.45), is the following:

$$L_2 = \prod_{t=1}^{T} \prod_{i=1}^{n_t} \prod_{j=1}^{J} \left\{ \begin{array}{l} P_j^N - \left[\rho_j^r P_j^N \right] I_{\left(y_{ti} < \underline{J}_t^P \right)} + \\ + \left[\begin{array}{c} -P_j^N \left(\sum_{\xi > j} \pi_{\xi - \underline{J}_t^P} \right) \rho^r + \\ + \left(\sum_{\zeta < j} P_\zeta^N \right) \pi_{j - \underline{J}_t^P} \rho^r \end{array} \right] I_{\left(\underline{J}_t^P \leq j \leq \bar{J}_t^P \right)} \end{array} \right\}^{I_{(y_{ti} = j)}} \tag{4.204}$$

[78]In this way, through the choice of τ, we can control the extent of distortions in the wage growth distributions caused by the presence of DRWR.

[79]The corresponding simplification of expression (4.66).

Details of implementation of parametric approach Adapting expression (4.17) to reflect the choice of Normality, the PDF of the 'actual' (rigidity-contaminated) wage growth rates will be given by

$$f_t\left(\dot{w}_{ti}\right)=\left\{\begin{array}{l}\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}_{ti}-\mu_{Nt}}{\sigma_N}\right)- \\ -\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}_{ti}-\mu_{Nt}}{\sigma_N}\right)\left[1-\Phi\left(-\frac{\dot{w}_{ti}-\mu_{\dot{P}t}}{\sigma_{\dot{P}}}\right)\right]\rho^r+ \\ +\Phi\left(-\frac{\dot{w}_{ti}-\mu_{Nt}}{\sigma_N}\right)\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}_{ti}-\mu_{\dot{P}t}}{\sigma_{\dot{P}}}\right)\rho^r\end{array}\right. \tag{4.205}$$

and the likelihood function for the simulated samples of independently drawn DRWR-contaminated wage growth rates by

$$\begin{aligned}L_1&=\prod_{t=1}^{T}\prod_{i=1}^{n_t}f_t\left(\dot{w}_{ti}\right)\\&=\prod_{t=1}^{T}\prod_{i=1}^{n_t}\left\{\begin{array}{l}\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}_{ti}-\mu_{Nt}}{\sigma_N}\right)- \\ -\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}_{ti}-\mu_{Nt}}{\sigma_N}\right)\left[1-\Phi\left(-\frac{\dot{w}_{ti}-\mu_{\dot{P}t}}{\sigma_{\dot{P}}}\right)\right]\rho^r+ \\ +\Phi\left(-\frac{\dot{w}_{ti}-\mu_{Nt}}{\sigma_N}\right)\frac{1}{\sigma_N}\phi\left(\frac{\dot{w}_{ti}-\mu_{\dot{P}t}}{\sigma_{\dot{P}}}\right)\rho^r\end{array}\right\}\end{aligned} \tag{4.206}$$

Experiment design

The parameters left to vary for the design of the simulation exercises are ρ^r and $\sigma_{\dot{P}}$, the value of the location parameter λ, as well as the values of T and n_t. In Table 4.8 we have collected the values taken by these parameters for the three sets of simulation experiments executed, with the exception that, instead of reporting $\sigma_{\dot{P}}$, we report the implied number of bins that contain the values of the support of the AID, which we denote by $\#q$.

First we consider the 'base case', which is identical to the 'base' case under DNWR, examined the previous section; ρ^r is set equal to 0.5, the number of samples from each heterogeneity group is set according to the rule given in (4.193), and the standardisation location parameter is set equal to the notional median. In addition, $\#q$ is set equal to 3, corresponding to $\sigma_{\dot{P}}=0.5$.

Then we consider three sets of variations relative to the base case: In Variation 1 we deviate from the base case and allow ρ^r to take a 'low' value ($=0.2$) for a subset of experiments and a 'high' value ($=0.8$) for another. In Variation 2 we change $\#q$ to 5, which corresponds to $\sigma_{\dot{P}}=0.95$. Finally, in Variation 3 we set the location parameter equal to sample values of certain quantiles of the 'actual' WGD; the sample median, and the 60'th and 90'th sample percentiles. The sample median could be considered as the 'default' choice, in the sense that this has been the typical choice of location parameter among the studies in the literature that have dealt with the testing for DNWR only. Together with the other two choices, that lie to the right of the sample median, these will allow us to examine the importance of the choice of standardisation parameter, especially in the presence of DRWR.[80]

[80]We recall from the discussion in Section 4.4.4 that this should be made according to the position of the AID relative to the notional WGD, which often implies that the standardisation parameter has to be an

For each of the above designs we perform experiments where the number of samples T can take values from the set $\{5, 10, 20, 40, 80\}$, and the sample size n_t from the set $\{63, 125, 250, 500\}$. Each experiment is replicated 50 times, except where indicated in the results' tables.

Results

The base case and Variation 1 The results for the 'base' case and Variation 1 (see Tables 4.9 and 4.10) are qualitatively similar to the corresponding results described for the case of DNWR; that is, the estimator of the rigidity parameter - ρ^r in this case - appears to exhibit a small positive bias, and to be consistent. Furthermore, the size of the variance of the estimator of the rigidity parameter seems to be inversely related to the parameter's true value. The only difference from the case of DNWR that is, probably, worthy of mentioning is that the size of the bias in this case appears to be smaller.

Variation 2: vary spread of AID Increasing the spread of the AID seems to increase the variability of the estimator, that is, to make it less precise (see Table 4.11). This could be explained by the fact that, other things equal, the increase in the spread of the AID means that the probability mass that is shifted from the left is now allocated to a bigger range of values (the support of AID), therefore the size of extra mass allocated at each of these points - i.e. the size of the distortion - must be smaller. It is also interesting to note that, for some values of (T, n_t), the semiparametric estimator seems to outperform the parametric estimator when judged with respect to their corresponding values of the RMSE. This appears to be primarily due to the semiparametric estimator having a smaller variance.

Variation 3: vary standardisation parameter The results in Table 4.12 suggest that the choice of location parameter is critical for the properties of the estimator of ρ^r:

Firstly, the comparison of the results for $\hat{q}_{50}$ and those for q_{50} (in Table 4.9), shows that using an estimate of a location parameter of the notional WGD rather than its true value as a standardisation parameter can result to estimators with distinctly different properties: in this case, the estimator based on $\hat{q}_{50}$ exhibits considerable negative bias, whereas the mean of the estimator based on q_{50} appeared to be very close the true value, albeit with a small positive bias.

Secondly, the comparison of the results for $\hat{q}_{50}$, $\hat{q}_{60}$ and $\hat{q}_{90}$ seem to support the validity of condition (4.164) as a guide to the choice a standardisation parameter that is an *estimate*

estimate of a quantile of a higher order ($= \alpha$) than the median. Expression (4.164), in particular, gives the condition that must be satisfied by the chosen standardisation parameter that is an estimate of a population quantile, so that the log-likelihood function maximized is valid asymptotically. Given the nature of the DGP underlying the simulated 'actual' wage growth rates, only $\hat{q}_{90}$ marginally satisfies this condition, which implies $\min_t \left(\Pr \left(\dot{P}^e_{ti} < F_t^{-1}(\alpha) \right) \right) = 1$. In particular, we have $\min_t \left(\Pr \left(\dot{P}^e_{ti} < F_t^{-1}(0.9) \right) \right) = 0.993$, while $\min_t \left(\Pr \left(\dot{P}^e_{ti} < F_t^{-1}(0.5) \right) \right) = \min_t \left(\Pr \left(\dot{P}^e_{ti} < F_t^{-1}(0.6) \right) \right) = 0$

of a location parameter of the notional WGD that is based on actual wage growth data. As it can be seen, only the results based on $\hat{q}_{90}$ compares satisfactorily to the results obtained for the base case. On the other hand, the use of $\hat{q}_{60}$ appears to lead to negative bias, as was the result from the use of $\hat{q}_{50}$.

4.5.3 Data contaminated by both DNWR and DRWR

Basic setup

The DGP In this case, the DGP for the simulated 'actual' wage growth distributions is described by the model presented in Section 4.2, where $p^n, p^r \in (0,1)$.

For practical purposes we express the relationship between the generated 'actual' ($\dot{w}_{ti}$) and the corresponding 'notional' ($\dot{w}_{ti}^N$) wage growth rates in the following way

$$\dot{w}_{ti} = (1 - d_{ti}^n - d_{ti}^r) \times \dot{w}_{ti}^N + d_{ti}^n \times \dot{w}_{ti}^n + d_{ti}^r \times \dot{w}_{ti}^r \tag{4.207}$$

where $\dot{w}_{ti}^n$ and $\dot{w}_{ti}^r$ relate to $\dot{w}_{ti}^N$ in the way described by equations (4.181)-(4.182) and (4.194)-(4.195), respectively, and

$$d_{ti}^n = \begin{cases} 1 & , \quad d_{ti}^{*^n} \leq p^n \\ 0 & , \quad d_{ti}^{*^n} > p^n \end{cases} \quad , \quad d_{ti}^{*^n} \overset{IID}{\sim} U(0,1) \tag{4.208}$$

$$d_{ti}^r = \begin{cases} 1 & , \quad d_{ti}^{*^r} \leq p^r \\ 0 & , \quad d_{ti}^{*^r} > p^r \end{cases} \quad , \quad d_{ti}^{*^r} \overset{IID}{\sim} U(0,1) \tag{4.209}$$

The notional wage growth rates are again sampled from the Normal distribution

$$\dot{w}_{ti}^N = \mu_{Nt} + \epsilon_{ti}^N \tag{4.210}$$

$$\epsilon_{ti}^N \overset{IID}{\sim} N\left(0, \sigma_N^2\right) \quad , \quad \sigma_N \simeq 4.1 \tag{4.211}$$

where the variance is the same across samples, but the mean is allowed to take its values according to the values of parameter τ

$$\mu_{Nt} = \begin{cases} \mu_H \; = 18 & , \quad t \leq \tau \\ \mu_L \overset{IID}{\sim} U(0,4) & , \quad \tau < t \leq T \end{cases} \tag{4.212}$$

The values of the anticipated inflation rates are also sampled from the Normal distribution

$$\dot{P}_{ti}^e = \mu_{\dot{P}t} + \epsilon_{ti}^P \tag{4.213}$$

$$\epsilon_{ti}^P \overset{IID}{\sim} N\left(0, \sigma_{\dot{P}}^2\right) \quad , \quad \sigma_{\dot{P}} \simeq 0.5 \tag{4.214}$$

where the standard deviation is fixed across t, but the mean takes values according to the values of parameter τ

$$\mu_{\dot{P}t} = \begin{cases} 14 & , \quad t \leq \tau \\ \mu_L + 4 & , \quad \tau < t \leq T \end{cases} \tag{4.215}$$

With the above assumptions, the shape of the 'notional' WGD is the same across t and only its location is allowed to change. The same is also true for the AID. Furthermore, the location of the 'notional' WGD is such that the presence of DNWR can only affect the shape of the 'actual' WGD for $t > \tau$. For the same values of t, the position of the centre of the AID is fixed relative to centre of the notional WGD, and located to its right. For $t \leq \tau$, the relative position is also fixed, but located to its left.[81] Consequently, the T simulated samples of 'actual' wage growth rates are generated from a collection of distributions which are distorted in different ways by the two types of rigidity.[82]

For all the simulation exercises considered in this section, the DGP of the simulated 'actual' wage growth rates is the same across experiments. The values of the parameters that define this DGP, in addition to those whose values were defined above, are the following;

$$p^n = 0.4 \quad , \quad \rho^n = 0.5 \quad \Rightarrow \rho^{nn} = 0.2 \tag{4.216}$$

$$p^r = 0.4 \quad , \quad \rho^r = 0.5 \quad \Rightarrow \rho^{rr} = 0.2 \tag{4.217}$$

and τ is determined according to the rule given in (4.193).

Details of implementation of semiparametric and parametric approach The details regarding the partition of the support of the notional distribution for the implementation of the 'discretisation' stage of the semiparametric approach are the same as in the case of DRWR. Furthermore, the location parameter λ_t is left as a simulation parameter. In this case, the log-likelihood function for the sample of the discrete random variables

$$\{y_{ti}\}_{t=1,\ldots,T,\; i=1,\ldots,n_t} \tag{4.218}$$

is given by (4.76).

On the other hand, the details of the implementation of the parametric approach under Normality have already been discussed in Section 4.3. In particular, the relevant likelihood function for the simulated 'actual' wage growth rates under both types of downward rigidity is given by (4.26). We recall that both the parametric and semiparametric approaches can only identify the total-rigidity parameters, i.e. ρ^{nn} and ρ^{rr}.

Experiment design

The simulation work undertaken here concentrates on providing information about the implications of the choice of standardisation parameter for the properties of the estimators of the total-rigidity parameters, ρ^{nn} and ρ^{rr}. For this reason, the main parameter that varies across simulation designs is the value of the standardisation parameter used for the standardisation. As before, we also allow for variation in T and n_t.

In Table 4.13 we have collected the values taken by these parameters[83] for two sets

[81] The relative position of the notional WGD and AID is the same as in the case of DRWR for $t \leq \tau$ and $t > \tau$.

[82] This ensures the identification of the rigidity parameters.

[83] For easy reference, we have also included the values of the remaining parameters that were treated as simulation parameters in the previous two cases.

of simulation exercises. The 'base case' is defined in the same way as in the case where only DRWR was present. Also Variation 1 resembles the set of experiments referred to as 'Variation 3', also in that case. For each of the above designs we perform experiments where the number of samples T takes values $\{10, 20\}$, and for each of these values, the sample size n_t takes values $\{63, 125, 250, 500\}$. Each experiment is replicated 50 times.

Results

The results obtained here verify the conclusions derived for the previous two cases.

In particular, we find that when the standardisation parameter is the population value of a location parameter of the notional WGD, as it is for the 'base' case (see Table 4.14), the (total-)rigidity parameter estimators appear to be consistent, with mean close to the true value - even for small sample sizes - but positively biased.

On the other hand, given that the presence of DRWR results to distortions in the shape of the actual WGD that extends beyond the centre of the corresponding notional WGD, we verify that standardisation that is based on the sample median leads to biased estimates of the rigidity parameters (see Table 4.15). Specifically, this seems to affect mainly the rigidity parameter associated with DRWR (ρ^{rr}), which appears to be negatively biased.

The same conclusion is reached for the case where standardisation is based on the sample 80'th percentile (see Table 4.16); this is expected since $\hat{q}_{80,t}$ does not satisfy condition (4.164).[84]

On the other hand, when standardisation is based on the sample 90'th percentile (see Table 4.17) the results concerning the rigidity parameter estimators for both types of rigidity are comparable to those obtained under the base case - where standardisation was based on the true value of a location parameter of the notional WGD.

4.5.4 Computational details

The simulations described in the previous sub-sections were performed in STATA. In particular, a custom-made code was written in order to simulate the data,[85] and the maximisation of the likelihood functions corresponding to the semiparametric and parametric estimators was performed with the `ml` command, using the `lf` method that implemented the Newton-Raphson algorithm.

In order to provide some feel for the nature of the calculations involved for the implementation of the semiparametric and parametric estimators, in Table 4.18 we report two statistics regarding the convergence of the maximisation algorithm for the three types of DGPs involved in the simulations, that were discussed in Sections 4.5.1-4.5.3. These statistics are the averages, calculated across the relevant number of replications of each experiment, of the time taken to convergence ('avg. duration'), and the number of iterations of the algorithm to convergence ('avg. #iterations'). We only report the statistics for

[84] In fact we have $\min_t\left(\Pr\left(\dot{P}^e_{ti} < F_t^{-1}(0.8)\right)\right) = 0.13$.

[85] That is, instead of using the in-built `simulate` command.

selected experiments[86] from the corresponding 'base' cases, as similar values apply for the experiments of the same characteristic in the other designs. The experiments considered are those involving the smallest and biggest sample sizes for each type of DGP, as well as the cases of $T = 10/n_t = 63$ and $T = 20/n_t = 500$ for all three DGP types, for comparison purposes.[87]

4.6 Conclusion

In this paper we present a semiparametric approach to estimating the rigidity parameters of a model with proportional DNWR and DRWR, similar to the one whose parametric estimation is discussed by Goette et al. (2007). Our approach could be implemented with other models of downward wage rigidity as well.

It is based on the principles that underly the location-histogram approach of Kahn (1997), and we show that the two approaches are asymptotically equivalent.

Our approach is developed within a formal framework, where the underlying assumptions are clearly stated. We provide some theoretical investigation of its properties, and also discuss sufficient conditions for identification.

Complementary to this analysis, we also use execute simulations to study some of its properties, as well as compare its performance to the parametric approach. The results point out, in particular, to the importance of the choice of standardisation parameter, especially when DRWR is thought to be present. Also, its performance is judged positively against the performance of the parametric approach.

[86]Determined by the values of T and n_t.

[87]These are the values of T and n_t for the experiments involving the smallest and biggest sample sizes in the case where both DNWR and DRWR are allowed to be present. The duration values reported in Table 4.18 depend on the technical specifications of the computer used for the calculations as well as the particular way the maximisation calculations are coded up, especially those for the ('non-standard') semiparametric estimator. Although great effort was made to write the code for the implementation of this estimator as efficiently as possible, there might be room for improvement that could potentially lead to a reduction in the apparent differences between in the time required for the implementation of the two estimators.

Base Case	
Parameter	**Simulation values**
ρ^n	0.5
(T, τ)	$(5,2), (10,5), (20,10), (40,20), (80,40)$
n_t	$63, 125, 250, 500$
λ	m_t^N
[See Table 4.2 for results.]	

Variation 1: Vary ρ^n	
Parameter	**Simulation values**
ρ^n	$0.2, 0.8$
(T, τ)	$(5,2), (10,5), (20,10), (40,20), (80,40)$
n_t	$63, 125, 250, 500$
λ	m_t^N
[See Tables 4.3 and 4.4 for results.]	

Variation 2: Vary τ	
Parameter	**Simulation values**
ρ^n	0.5
T	$5, 10, 20, 40, 80$
τ	$1, T$
n_t	$63, 125, 250, 500$
λ	m_t^N
[See Tables 4.5 and 4.6 for results.]	

Variation 3: Vary λ	
Parameter	**Simulation values**
ρ^n	0.5
(T, τ)	$(5,2), (10,5), (20,10), (40,20), (80,40)$
n_t	$63, 125, 250, 500$
λ	$\hat{m}_t$
[See Table 4.7 for results.]	

Table 4.1: Simulation design: DNWR.

			Semiparametric				Parametric		
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$	*mean*	*s.d.*	*RMSE*
5	63	315	0.5127	0.0999	0.1002	304	0.5023	0.0786	0.0782
5	125	625	0.5069	0.0749	0.0748	605	0.4931	0.0612	0.0613
5	250	1250	0.5151	0.0531	0.0550	1208	0.5016	0.0369	0.0368
5	500	2500	0.5134	0.0380	0.0401	2416	0.5000	0.0307	0.0305
10	63	630	0.5189	0.0705	0.0726	609	0.5092	0.0545	0.0550
10	125	1250	0.5122	0.0465	0.0478	1211	0.4982	0.0357	0.0356
10	250	2500	0.5128	0.0281	0.0308	2419	0.5004	0.0251	0.0249
10	500	5000	0.5153	0.0222	0.0268	4835	0.5002	0.0194	0.0193
20	63	1260	0.5166	0.0446	0.0474	1219	0.5023	0.0342	0.0341
20	125	2500	0.5089	0.0316	0.0327	2420	0.4955	0.0257	0.0260
20	250	5000	0.5154	0.0233	0.0279	4836	0.5018	0.0190	0.0190
20	500	10000	0.5133	0.0148	0.0199	9671	0.5002	0.0130	0.0129
40	63	2520	0.5194	0.0316	0.0369	2438	0.5049	0.0241	0.0245
40	125	5000	0.5091	0.0222	0.0239	4839	0.4956	0.0190	0.0194
40	250	10000	0.5138	0.0147	0.0201	9673	0.5006	0.0132	0.0131
40	500	20000	0.5141	0.0098	0.0171	19345	0.5009	0.0091	0.0091
80	63	5040	0.5179	0.0229	0.0290	4876	0.5050	0.0170	0.0176
80	125	10000	0.5114	0.0158	0.0195	9676	0.4981	0.0139	0.0140
80	250	20000	0.5129	0.0105	0.0166	19348	0.5007	0.0093	0.0093
80	500	40000	0.5133	0.0074	0.0152	38693	0.5006	0.0061	0.0061

$\rho^n = 0.5$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = m_t^N$

Table 4.2: Results: DNWR, base case.

			Semiparametric				Parametric		
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$	*mean*	*s.d.*	*RMSE*
5	63	315	0.2187	0.0999	0.1011	303	0.2013	0.0698	0.0694
5	125	625	0.2137	0.0789	0.0797	603	0.2052	0.0453	0.0454
5	250	1250	0.2098	0.0592	0.0597	1205	0.2035	0.0314	0.0315
5	500	2500	0.2117	0.0362	0.0379	2411	0.2013	0.0233	0.0232
10	63	630	0.2165	0.0680	0.0696	607	0.2076	0.0359	0.0366
10	125	1250	0.2090	0.0516	0.0521	1208	0.2027	0.0301	0.0301
10	250	2500	0.2078	0.0307	0.0315	2412	0.2025	0.0217	0.0218
10	500	5000	0.2105	0.0222	0.0245	4821	0.2007	0.0145	0.0145
20	63	1260	0.2154	0.0466	0.0489	1216	0.2046	0.0294	0.0296
20	125	2500	0.2093	0.0349	0.0359	2414	0.2011	0.0211	0.0210
20	250	5000	0.2105	0.0252	0.0272	4822	0.2036	0.0154	0.0157
20	500	10000	0.2088	0.0156	0.0178	9643	0.2015	0.0106	0.0107
40	63	2520	0.2149	0.0296	0.0330	2431	0.2054	0.0198	0.0204
40	125	5000	0.2077	0.0253	0.0263	4826	0.2005	0.0170	0.0169
40	250	10000	0.2080	0.0175	0.0192	9646	0.2018	0.0111	0.0112
40	500	20000	0.2087	0.0111	0.0140	19289	0.2011	0.0075	0.0075
80	63	5040	0.2119	0.0223	0.0252	4862	0.2043	0.0140	0.0146
80	125	10000	0.2090	0.0161	0.0184	9649	0.2013	0.0100	0.0101
80	250	20000	0.2072	0.0111	0.0132	19292	0.2010	0.0078	0.0078
80	500	40000	0.2079	0.0082	0.0114	38580	0.2011	0.0051	0.0052

$\rho^n = 0.2$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = m_t^N$

Table 4.3: Results: DNWR, change to $\rho^n = 0.2$ (relative to base case).

			Semiparametric				Parametric		
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$	*mean*	*s.d.*	*RMSE*
5	63	315	0.8172	0.0641	0.0661	305	0.8047	0.0601	0.0600
5	125	625	0.8032	0.0601	0.0599	606	0.7929	0.0547	0.0549
5	250	1250	0.8105	0.0359	0.0372	1211	0.8014	0.0334	0.0333
5	500	2500	0.8129	0.0262	0.0291	2422	0.8026	0.0228	0.0228
10	63	630	0.8184	0.0426	0.0463	611	0.8105	0.0395	0.0406
10	125	1250	0.8093	0.0343	0.0354	1214	0.7992	0.0310	0.0308
10	250	2500	0.8049	0.0649	0.0647	2412	0.8031	0.0181	0.0183
10	500	5000	0.8130	0.0173	0.0215	4849	0.8025	0.0150	0.0151
20	63	1260	0.8135	0.0323	0.0348	1222	0.8046	0.0289	0.0291
20	125	2500	0.8057	0.0216	0.0223	2427	0.7968	0.0196	0.0197
20	250	5000	0.8094	0.0155	0.0181	4850	0.8014	0.0142	0.0142
20	500	10000	0.8102	0.0116	0.0154	9699	0.8011	0.0097	0.0097
40	63	2520	0.8138	0.0197	0.0240	2444	0.8046	0.0185	0.0190
40	125	5000	0.8059	0.0142	0.0153	4853	0.7975	0.0130	0.0131
40	250	10000	0.8084	0.0106	0.0135	9702	0.8001	0.0100	0.0100
40	500	20000	0.8102	0.0072	0.0125	19401	0.8011	0.0061	0.0062
80	63	5040	0.8140	0.0147	0.0203	4890	0.8049	0.0135	0.0143
80	125	10000	0.8076	0.0106	0.0130	9704	0.7992	0.0099	0.0099
80	250	20000	0.8077	0.0074	0.0107	19403	0.7998	0.0067	0.0067
80	500	40000	0.8090	0.0054	0.0105	38805	0.8003	0.0049	0.0049

$\rho^n = 0.8$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = m_t^N$

Table 4.4: Results: DNWR, change to $\rho^n = 0.8$ (relative to base case).

			Semiparametric				Parametric		
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$	*mean*	*s.d.*	*RMSE*
5	63	315	0.5001	0.1452	0.1445	304	0.4914	0.1071	0.1069
5	125	625	0.5070	0.1204	0.1200	603	0.4928	0.0900	0.0898
5	250	1250	0.5179	0.0787	0.0803	1206	0.5046	0.0584	0.0583
5	500	2500	0.5167	0.0562	0.0584	2412	0.5056	0.0417	0.0418
10	63	630	0.5038	0.1455	0.1448	607	0.4956	0.1077	0.1072
10	125	1250	0.5084	0.1189	0.1186	1206	0.4959	0.0898	0.0894
10	250	2500	0.5168	0.0782	0.0796	2409	0.5092	0.0577	0.0581
10	500	5000	0.5166	0.0552	0.0573	4817	0.5094	0.0412	0.0420
20	63	1260	0.5053	0.1405	0.1399	1214	0.5010	0.1087	0.1082
20	125	2500	0.5092	0.1141	0.1139	2411	0.5040	0.0879	0.0876
20	250	5000	0.5172	0.0794	0.0809	4815	0.5169	0.0563	0.0585
20	500	10000	0.5168	0.0550	0.0572	9629	0.5180	0.0410	0.0446
40	63	2520	0.5065	0.1397	0.1392	2427	0.5159	0.1030	0.1037
40	125	5000	0.5105	0.1125	0.1124	4818	0.5199	0.0875	0.0893
40	250	10000	0.5174	0.0793	0.0808	9629	0.5324	0.0551	0.0637
40	500	20000	0.5175	0.0544	0.0569	19256	0.5337	0.0406	0.0526
80	63	5040	0.5070	0.1393	0.1388	4854	0.5414	0.1032	0.1107
80	125	10000	0.5113	0.1108	0.1108	9632	0.5469	0.0839	0.0958
80	250	20000	0.5179	0.0781	0.0798	19257	0.5610	0.0547	0.0817
80	500	40000	0.5173	0.0544	0.0569	38511	0.5616	0.0421	0.0745

$\rho^n = 0.5$
$\tau = 1$
$\lambda = m_t^N$

Table 4.5: Results: DNWR, change to $\tau = 1$ (relative to base case).

			Semiparametric				Parametric		
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$	*mean*	*s.d.*	*RMSE*
5	63	315	0.5159	0.0850	0.0861	306	0.5081	0.0538	0.0541
5	125	625	0.5150	0.0501	0.0520	608	0.4984	0.0366	0.0364
5	250	1250	0.5140	0.0309	0.0338	1215	0.4995	0.0252	0.0250
5	500	2500	0.5170	0.0242	0.0295	2430	0.4995	0.0193	0.0192
10	63	630	0.5170	0.0487	0.0514	612	0.5018	0.0341	0.0340
10	125	1250	0.5099	0.0340	0.0353	1216	0.4948	0.0258	0.0262
10	250	2500	0.5158	0.0240	0.0286	2431	0.5011	0.0190	0.0190
10	500	5000	0.5152	0.0156	0.0217	4859	0.4994	0.0130	0.0129
20	63	1260	0.5201	0.0315	0.0372	1225	0.5041	0.0243	0.0245
20	125	2500	0.5097	0.0238	0.0256	2432	0.4947	0.0189	0.0195
20	250	5000	0.5145	0.0149	0.0208	4859	0.4998	0.0132	0.0132
20	500	10000	0.5154	0.0101	0.0184	9719	0.5001	0.0091	0.0091
40	63	2520	0.5192	0.0230	0.0299	2449	0.5043	0.0171	0.0175
40	125	5000	0.5127	0.0165	0.0208	4862	0.4974	0.0140	0.0141
40	250	10000	0.5142	0.0109	0.0178	9720	0.4999	0.0093	0.0092
40	500	20000	0.5148	0.0074	0.0165	19438	0.4999	0.0060	0.0060
80	63	5040	0.5176	0.0144	0.0228	4899	0.5024	0.0112	0.0114
80	125	10000	0.5146	0.0119	0.0188	9722	0.4990	0.0089	0.0089
80	250	20000	0.5148	0.0074	0.0166	19441	0.5004	0.0057	0.0057
80	500	40000	0.5147	0.0058	0.0158	38878	0.5003	0.0044	0.0044

$\rho^n = 0.5$
$\tau = T$
$\lambda = m_t^N$

Table 4.6: Results: DNWR, change to $\tau = T$ (relative to base case).

			Semiparametric			
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$
5	63	315	0.5070	0.1060	0.1057	304
5	125	625	0.5052	0.0754	0.0752	604
5	250	1250	0.5152	0.0472	0.0494	1208
5	500	2500	0.5119	0.0411	0.0426	2416
10	63	630	0.5170	0.0656	0.0674	609
10	125	1250	0.5096	0.0493	0.0500	1210
10	250	2500	0.5097	0.0310	0.0323	2419
10	500	5000	0.5139	0.0236	0.0273	4834
20	63	1260	0.5105	0.0456	0.0466	1220
20	125	2500	0.5069	0.0330	0.0336	2420
20	250	5000	0.5151	0.0230	0.0275	4835
20	500	10000	0.5125	0.0163	0.0204	9670
40	63	2520	0.5170	0.0328	0.0368	2438
40	125	5000	0.5065	0.0234	0.0242	4839
40	250	10000	0.5139	0.0158	0.0209	9672
40	500	20000	0.5136	0.0105	0.0171	19343
80	63	5040	0.5150	0.0233	0.0276	4875
80	125	10000	0.5099	0.0164	0.0191	9674
80	250	20000	0.5130	0.0110	0.0170	19345
80	500	40000	0.5128	0.0075	0.0148	38691

$\rho^n = 0.5$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = \hat{m}_t$

Table 4.7: Results: DNWR, change to $\lambda = \hat{m}_t$ (relative to base case).

Base Case	
Parameter	**Simulation values**
ρ^r	0.5
$\#q$	3
(T, τ)	$(5,2)\,,(10,5)\,,(20,10)\,,(40,20)\,,(80,40)$
n_t	$63, 125, 250, 500$
λ	m_t^N
[See Table 4.9 for results.]	
NB: When $T = 80$, then $n_t \in \{63, 125, 250\}$	

Variation 1: Vary ρ^r	
Parameter	**Simulation values**
ρ^r	$0.2, 0.8$
$\#q$	3
(T, τ)	$(5,2)\,,(10,5)\,,(20,10)\,,(40,20)\,,(80,40)$
n_t	$63, 125, 250, 500$
λ	m_t^N
[See Table 4.10 for results.]	
NB: When $T = 80$, then $n_t \in \{63, 125, 250\}$	

Variation 2: Vary $\#q$	
Parameter	**Simulation values**
ρ^r	0.2
$\#q$	5
(T, τ)	$(5,2)\,,(10,5)\,,(20,10)\,,(40,20)$
n_t	$63, 125, 250, 500$
λ	m_t^N
[See Table 4.11 for results.]	

Variation 3: Vary λ	
Parameter	**Simulation values**
ρ^r	0.5
$\#q$	3
(T, τ)	$(5,2)\,,(10,5)\,,(20,10)\,,(40,20)$
n_t	$63, 125, 250, 500$
λ	$\hat{m}_t (= \hat{q}_{50}),\ \hat{q}_{60},\ \hat{q}_{90}$
[See Table 4.12 for results.]	

Table 4.8: Simulation design: DRWR.

T	n_t	$T \times n_t$	Semiparametric				Parametric		
			$mean$	$s.d.$	$RMSE$	$\overline{T \times n_t}$	$mean$	$s.d.$	$RMSE$
5	63	315	0.4919	0.0456	0.0458	305	0.4925	0.0812	0.0807
5	125	625	0.4994	0.0322	0.0318	606	0.4917	0.0265	0.0275
5	250	1250	0.5091	0.0238	0.0253	1211	0.5028	0.0204	0.0204
5	500	2500	0.5050	0.0180	0.0185	2423	0.4958	0.0181	0.0184
10	63	630	0.5005	0.0413	0.0408	610	0.4990	0.0338	0.0334
10	125	1250	0.5116	0.0266	0.0288	1212	0.5017	0.0236	0.0234
10	250	2500	0.5041	0.0159	0.0162	2423	0.4971	0.0160	0.0161
10	500	5000	0.5083	0.0166	0.0185	4844	0.5008	0.0145	0.0144
20	63	1260	0.5068	0.0237	0.0245	1222	0.5057	0.0212	0.0217
20	125	2500	0.5064	0.0181	0.0190	2425	0.5026	0.0179	0.0179
20	250	5000	0.5079	0.0154	0.0171	4846	0.5000	0.0119	0.0117
20	500	10000	0.5057	0.0077	0.0095	9691	0.4998	0.0077	0.0076
40	63	2520	0.5062	0.0197	0.0204	2444	0.5026	0.0163	0.0163
40	125	5000	0.5066	0.0148	0.0161	4850	0.5009	0.0134	0.0133
40	250	10000	0.5062	0.0095	0.0112	9695	0.5000	0.0076	0.0075
40	500	20000	0.5064	0.0061	0.0088	19389	0.5015	0.0051	0.0053
80	63	5040	0.5088	0.0109	0.0139	4886	0.5017	0.0105	0.0105
80	125	10000	0.5063	0.0087	0.0107	9697	0.4886	0.0710	0.0712
80	250	20000	0.5073	0.0075	0.0104	19388	0.5009	0.0068	0.0067
$\rho^r = 0.5$									
$\#q = 3$									
$\tau = T/2$ if T even, $(T-1)/2$ if T odd									
$\lambda = m_t^N$									

Table 4.9: Results: DRWR, base case.

T	n_t	$T \times n_t$	Semiparametric				Semiparametric			
			$mean$	$s.d.$	$RMSE$	$\overline{T \times n_t}$	$mean$	$s.d.$	$RMSE$	$\overline{T \times n_t}$
5	63	315	0.2002	0.0504	0.0499	303	0.7933	0.0344	0.0347	306
5	125	625	0.2056	0.0355	0.0356	602	0.8053	0.0194	0.0199	609
5	250	1250	0.2117	0.0304	0.0322	1204	0.8072	0.0160	0.0174	1218
5	500	2500	0.2052	0.0173	0.0179	2409	0.8020	0.0123	0.0124	2437
10	63	630	0.2068	0.0425	0.0426	606	0.7981	0.0277	0.0275	614
10	125	1250	0.2134	0.0251	0.0282	1206	0.8015	0.0193	0.0192	1219
10	250	2500	0.2052	0.0184	0.0189	2409	0.8028	0.0136	0.0137	2437
10	500	5000	0.2066	0.0146	0.0159	4815	0.8053	0.0113	0.0124	4872
20	63	1260	0.2078	0.0303	0.0310	1215	0.8071	0.0186	0.0198	1228
20	125	2500	0.2037	0.0179	0.0181	2412	0.8038	0.0154	0.0157	2439
20	250	5000	0.2067	0.0145	0.0159	4818	0.8032	0.0082	0.0088	4875
20	500	10000	0.2025	0.0101	0.0103	9635	0.8029	0.0067	0.0072	9749
40	63	2520	0.2099	0.0171	0.0196	2430	0.8051	0.0118	0.0127	2457
40	125	5000	0.2059	0.0154	0.0164	4823	0.8045	0.0106	0.0114	4876
40	250	10000	0.2050	0.0092	0.0104	9639	0.8030	0.0062	0.0068	9749
40	500	20000	0.2047	0.0068	0.0082	19277	0.8049	0.0040	0.0063	19499
80	63	5040	0.2049	0.0133	0.0140	4859	0.8054	0.0075	0.0092	4914
80	125	10000	0.2063	0.0105	0.0122	9641	0.8048	0.0062	0.0078	9752
80	250	20000	0.2038	0.0078	0.0086	19279	0.8041	0.0054	0.0068	19497
ρ^r			0.2				0.8			
$\#q = 3$										
$\tau = T/2$ if T even, $(T-1)/2$ if T odd										
$\lambda = m_t^N$										

Table 4.10: Results: DRWR, change ρ^r (relative to base case).

			Semiparametric				Parametric		
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$	*mean*	*s.d.*	*RMSE*
5	63	315	0.5186	0.0523	0.0550	303	0.4940	0.0877	0.0870
5	125	625	0.5073	0.0399	0.0402	603	0.4855	0.0800	0.0805
5	250	1250	0.5083	0.0287	0.0296	1204	0.5069	0.0292	0.0297
5	500	2500	0.5036	0.0215	0.0216	2409	0.4947	0.0237	0.0241
10	63	630	0.5145	0.0431	0.0450	607	0.4955	0.0438	0.0436
10	125	1250	0.5099	0.0268	0.0283	1206	0.4978	0.0310	0.0308
10	250	2500	0.5023	0.0200	0.0199	2409	0.4959	0.0192	0.0194
10	500	5000	0.5070	0.0193	0.0204	4816	0.4987	0.0156	0.0155
20	63	1260	0.5073	0.0290	0.0296	1215	0.5048	0.0271	0.0273
20	125	2500	0.5092	0.0192	0.0212	2412	0.5064	0.0221	0.0228
20	250	5000	0.5092	0.0167	0.0189	4818	0.5024	0.0140	0.0140
20	500	10000	0.5070	0.0098	0.0120	9635	0.5021	0.0097	0.0098
40	63	2520	0.5056	0.0203	0.0208	2430	0.5035	0.0342	0.0340
40	125	5000	0.5066	0.0159	0.0171	4821	0.4997	0.0190	0.0189
40	250	10000	0.5070	0.0130	0.0147	9635	0.4907	0.0714	0.0713
40	500	20000	0.5072	0.0080	0.0107	19276	0.5013	0.0070	0.0070

$\rho^r = 0.5$
$\#q = 5$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = m_t^N$

Table 4.11: Results: DRWR, change to $\#q = 5$ (relative to base case).

			Semiparametric			Semiparametric			Semiparametric		
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	*mean*	*s.d.*	*RMSE*	*mean*	*s.d.*	*RMSE*
5	63	315	0.3942	0.0676	0.1252	0.4438	0.0521	0.0763	0.4918	0.0506	0.0507
5	125	625	0.3708	0.0602	0.1423	0.4421	0.0475	0.0746	0.4840	0.0374	0.0404
5	250	1250	0.3905	0.0620	0.1255	0.4521	0.0323	0.0576	0.4973	0.0285	0.0284
5	500	2500	0.3929	0.0578	0.1214	0.4444	0.0237	0.0604	0.4895	0.0224	0.0245
10	63	630	0.3999	0.0536	0.1133	0.4481	0.0442	0.0679	0.4901	0.0426	0.0433
10	125	1250	0.3946	0.0412	0.1130	0.4432	0.0325	0.0653	0.4999	0.0281	0.0278
10	250	2500	0.3939	0.0344	0.1114	0.4385	0.0257	0.0666	0.4901	0.0200	0.0222
10	500	5000	0.3935	0.0343	0.1117	0.4486	0.0213	0.0556	0.4979	0.0163	0.0162
20	63	1260	0.3904	0.0339	0.1146	0.4420	0.0299	0.0651	0.4948	0.0251	0.0254
20	125	2500	0.3857	0.0293	0.1179	0.4440	0.0214	0.0599	0.4931	0.0201	0.0211
20	250	5000	0.3971	0.0201	0.1048	0.4431	0.0157	0.0590	0.4938	0.0160	0.0170
20	500	10000	0.4046	0.0189	0.0972	0.4455	0.0119	0.0558	0.4936	0.0093	0.0112
40	63	2520	0.3908	0.0321	0.1137	0.4426	0.0265	0.0631	0.4942	0.0197	0.0203
40	125	5000	0.3886	0.0220	0.1135	0.4450	0.0144	0.0567	0.4951	0.0146	0.0153
40	250	10000	0.3978	0.0188	0.1039	0.4422	0.0123	0.0591	0.4948	0.0099	0.0111
40	500	20000	0.4099	0.0128	0.0910	0.4463	0.0099	0.0546	0.4946	0.0075	0.0092
λ				$\hat{m}_t^N$			$\hat{q}_{60}$			$\hat{q}_{90}$	

$\rho^r = 0.5$
$\#q = 3$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd

Table 4.12: Results: DRWR, change λ (relative to base case).

Base Case	
Parameter	**Simulation values**
(ρ^n, p^n)	$(0.5, 0.4)$
(ρ^r, p^r)	$(0.5, 0.4)$
$\#q$	3
(T, τ)	$(10, 5), (20, 10)$
n_t	$63, 125, 250, 500$
λ	m_t^N
[See Table 4.14 for results.]	

Variation 1: Vary λ	
Parameter	**Simulation values**
(ρ^n, p^n)	$(0.5, 0.4)$
(ρ^r, p^r)	$(0.5, 0.4)$
$\#q$	3
(T, τ)	$(10, 5), (20, 10)$
n_t	$63, 125, 250, 500$
λ	$\hat{m}_t (= \hat{q}_{50}),\ \hat{q}_{80},\ \hat{q}_{90}$
[See Tables 4.15-4.17 for results.]	

Table 4.13: Simulation design: DWR.

			Semiparametric, ρ^{nn}			Semiparametric, ρ^{rr}			
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$
10	63	630	0.2002	0.0537	0.0532	0.2146	0.0472	0.0489	608
10	125	1250	0.2029	0.0485	0.0481	0.2135	0.0251	0.0283	1206
10	250	2500	0.2100	0.0278	0.0293	0.2043	0.0238	0.0240	2413
10	500	5000	0.2022	0.0201	0.0200	0.2045	0.0137	0.0143	4827
20	63	1260	0.2111	0.0337	0.0352	0.2016	0.0296	0.0294	1217
20	125	2500	0.2085	0.0306	0.0315	0.2031	0.0187	0.0188	2416
20	250	5000	0.2011	0.0200	0.0198	0.2040	0.0143	0.0147	4830
20	500	10000	0.2072	0.0155	0.0170	0.2019	0.0115	0.0116	9652

$\rho^{nn} = 0.2$
$\rho^{rr} = 0.2$
$\#q = 3$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = m_t^N$

Table 4.14: Results: DWR, base case.

			Semiparametric, ρ^{nn}			Semiparametric, ρ^{rr}			
T	n_t	$T \times n_t$	*mean*	*s.d.*	*RMSE*	*mean*	*s.d.*	*RMSE*	$\overline{T \times n_t}$
10	63	630	0.1962	0.0667	0.0661	0.1648	0.0299	0.0460	603
10	125	1250	0.2084	0.0552	0.0552	0.1562	0.0213	0.0486	1198
10	250	2500	0.2208	0.0356	0.0409	0.1449	0.0229	0.0595	2398
10	500	5000	0.2116	0.0237	0.0262	0.1349	0.0131	0.0664	4801
20	63	1260	0.2201	0.0389	0.0435	0.1551	0.0212	0.0496	1208
20	125	2500	0.2229	0.0359	0.0423	0.1446	0.0169	0.0579	2399
20	250	5000	0.2068	0.0217	0.0225	0.1423	0.0139	0.0593	4797
20	500	10000	0.2131	0.0173	0.0215	0.1368	0.0093	0.0638	9600

$\rho^{nn} = 0.2$
$\rho^{rr} = 0.2$
$\#q = 3$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = \hat{m}_t$

Table 4.15: Results: DWR, change to $\lambda = \hat{m}_t$ (relative to base case).

T	n_t	$T \times n_t$	Semiparametric, ρ^{nn} mean	s.d.	RMSE	Semiparametric, ρ^{rr} mean	s.d.	RMSE	$\overline{T \times n_t}$
10	63	630	0.2078	0.0646	0.0645	0.1592	0.0397	0.0567	604
10	125	1250	0.1988	0.0548	0.0543	0.1448	0.0221	0.0594	1200
10	250	2500	0.2079	0.0353	0.0358	0.1546	0.0209	0.0499	2398
10	500	5000	0.2020	0.0222	0.0221	0.1602	0.0160	0.0428	4792
20	63	1260	0.2117	0.0414	0.0426	0.1482	0.0295	0.0595	1210
20	125	2500	0.2098	0.0319	0.0331	0.1522	0.0213	0.0522	2399
20	250	5000	0.2039	0.0234	0.0235	0.1558	0.0165	0.0471	4793
20	500	10000	0.2038	0.0172	0.0174	0.1613	0.0107	0.0402	9585

$\rho^{nn} = 0.2$
$\rho^{rr} = 0.2$
$\#q = 3$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = \hat{q}_{80}$

Table 4.16: Results: DWR, change to $\lambda = \hat{q}_{80}$ (relative to base case).

T	n_t	$T \times n_t$	Semiparametric, ρ^{nn} mean	s.d.	RMSE	Semiparametric, ρ^{rr} mean	s.d.	RMSE	$\overline{T \times n_t}$
10	63	630	0.2021	0.0501	0.0497	0.1979	0.0435	0.0431	603
10	125	1250	0.1981	0.0445	0.0441	0.1944	0.0251	0.0255	1196
10	250	2500	0.2151	0.0313	0.0345	0.1869	0.0235	0.0267	2389
10	500	5000	0.2065	0.0156	0.0168	0.1951	0.0175	0.0180	4772
20	63	1260	0.2100	0.0371	0.0381	0.1888	0.0257	0.0278	1206
20	125	2500	0.2089	0.0289	0.0299	0.1898	0.0220	0.0240	2392
20	250	5000	0.2088	0.0194	0.0212	0.1906	0.0181	0.0202	4776
20	500	10000	0.2072	0.0150	0.0165	0.1936	0.0113	0.0129	9552

$\rho^{nn} = 0.2$
$\rho^{rr} = 0.2$
$\#q = 3$
$\tau = T/2$ if T even, $(T-1)/2$ if T odd
$\lambda = \hat{q}_{90}$

Table 4.17: Results: DWR, change to $\lambda = \hat{q}_{90}$ (relative to base case).

DNWR					
		Semiparametric		Parametric	
T	n_t	avg. duration	avg. #iterations	avg. duration	avg. #iterations
5	63	16	3.95	2	5.00
10	63	25	3.93	2	5.05
20	500	350	4.10	10	5.03
80	500	2105	4.13	64	4.95
DRWR					
		Semiparametric		Parametric	
T	n_t	avg. duration	avg. #iterations	avg. duration	avg. #iterations
5	63	31	4.14	10	9.33
10	63	40	3.95	3	9.98
20	500	558	3.97	34	8.87
40	500	2425	4.00	100	9.51
DWR					
		Semiparametric		Parametric	
T	n_t	avg. duration	avg. #iterations	avg. duration	avg. #iterations
10	63	40	5.50	-	-
20	500	755	5.74	-	-

avg. duration ≡ average duration of time to convergence (measured in seconds)
avg. #iterations ≡ average number of algorithm iterations to convergence
NB: averages calculated across experiment replications

Table 4.18: Computational statistics ('base case' experiments).

Chapter 5

Conclusion

5.1 Summary

The work in this thesis is concerned with the development of a semiparametric approach for the joint testing of the hypotheses of Downward Nominal Wage Rigidity (DNWR) and Downward Real Wage Rigidity (DRWR) using micro-level data on nominal wage growth rates. Furthermore, with the implementation of the approach developed in order to test for the existence of both types of rigidity in the wages agreed through collective bargaining in the Canadian unionised sector.

The approach developed builds upon the methodology proposed by Kahn (1997) for the purpose of testing for DNWR only. Apart from extending this to allow the testing for the existence of DRWR as well, we make an effort to provide a formal framework within which the ideas originating in Kahn (1997) are developed into a test for downward wage rigidity, with the aim to provide theoretical support for its validity and to take a closer look at its properties. We also use simulations to compare the performance of the semiparametric approach to the performance of the parametric approach followed by Goette et al. (2007). Reassuringly, the semiparametric results were not very different from the parametric results, although, as one would expect given the additional structure in the parametric approach, the latter yielded better results overall.

The applied work reported here is based on data on wage growth rates from collective agreements from Canada, from the period 1976-1999. This particular data set has the advantage that it covers periods of diverse inflation experience, which is necessary for the identification of the distortions caused by the two types of rigidity. At the same time, it is free of measurement error, therefore no special treatment to address such issue is necessary. This application complements the work by Christofides and Leung (2003), who used a variant of the original Kahn (1997) approach to test for DNWR using this particular data set. The results obtained clearly suggest that DRWR exists over and above DNWR. Furthermore, fitting the model on data from subperiods characterised by different inflation levels has provided additional information regarding the patterns of the two types of rigidity as inflation changes. Specifically, during the high inflation period we find some evidence for the existence of DRWR while DNWR does not appear to be relevant. During the medium inflation period the results clearly suggest that both types of rigidity are present. This is also true for the low inflation period, except that, from the estimated size of the distortions attributed to DNWR, this appears to be the main type of rigidity that is relevant in years with low inflation. Finally, our results also show evidence of modest menu costs.

These findings are in agreement with the findings of earlier studies which investigated the presence of downward rigidity using the same data as those used here. In particular, evidence in favour of the presence of both types of rigidity has also been reported by Christofides and Li (2005), based on a different empirical strategy than the one followed here. They also found, as we have here, a negative correlation between the intensity of DNWR and the level of inflation. This latter finding is also supported by the results obtained by Christofides and Stengos (2002), Christofides and Stengos (2003), and Christofides and Leung (2003), whose analysis is restricted to the investigation of DNWR only.

Interestingly, similar results have also been obtained by studies which have looked for evidence for both types of rigidity in other countries, using data on wage growth rates of individuals. In particular, results for the UK, Germany, and Italy reported in Goette et al. (2007) indicate the presence of both types of downward rigidity as well as the change in their relative importance in the way discussed above, i.e. with DNWR becoming more important during periods of low inflation, and DRWR during periods of high inflation. Similarly, results obtained by the International Wage Flexibility Project (IWFP), summarised in Dickens et.al. (2007), indicate the presence of both types of rigidity in many of the countries considered in their study, and an increase in the fraction of workers affected by DNWR as inflation falls. As Goette et al. (2007) discuss in their paper, such findings may have important implications for monetary policy; as Akerlof et al. (1996) have shown from the study of a macro model that assumes this type of patterns for downward wage rigidity and inflation level, in such a case a long-run zero inflation monetary policy is suboptimal relative to one aiming for moderate inflation levels.

We conclude by making reference to an insight provided by our estimation results that has important implications for the validity of earlier studies in this literature that ignored the presence of DRWR. In particular, this suggests that failure to account for the presence of DRWR when measuring DNWR could lead to bias in the estimated size of DNWR, which we find that it could be either positive or negative. A similar result is also reported in Goette et al. (2007), who concluded that this bias should be positive.

5.2 Further research

The semiparametric approach discussed here, particularly in Chapter 4, could, potentially, be used as a complement to the parametric approach described in Goette et al. (2007). In particular, to provide robustness checks for the parametric results as, typically, theory does not provide information on the parametric family of the (rigidity-free) wage growth distribution. Having said that, for this to become feasible, further work is required in two specific directions, especially if this approach is to be used in conjunction with survey data:

Firstly, towards extending the semiparametric approach in order to handle the presence of measurement error in the reported wage growth rates. Possibly, this could be achieved with the introduction of a preliminary stage to 'correct' the observed data for measurement error, as was done in the IWFP.

Secondly, further study is required in order to provide better understanding on how to handle individual heterogeneity in the mean of the rigidity-free wage growth distribution, especially, on how to account for characteristics which are continuous in nature. Also, on how to account for heterogeneity in the extent of each type of rigidity within the corresponding sub-populations affected by each type (measured by ρ^n and ρ^r), as well as heterogeneity in the relative size these sub-populations (measured by p^n and p^r). It would be particularly interesting to attempt to exploit heterogeneity in order to separately identify ρ^n and p^n, as well as ρ^r and p^r, which is not feasible under the current model setup.

In addition to the above, further research is also warranted on the issue of optimal choice of bin with - in the discretisation stage, - in view of the fact that this also affects the identification of the amount of probability mass of the anticipated inflation distribution that is allocated across bins.

It might also be interesting to pursue extensions to the semiparametric model considered in Chapter 4 that go beyond the features characterising its richer parametric version described in Goette et al. (2007).

One particular direction might be to account for the presence of dependence in the observed individual wage growth rates, either within, or/and across years - or, more generally, heterogeneity groups. A first step towards this might to use simulations in order to investigate the effect of ignoring the presence of this feature in the data generating process. That is, simulate data assuming this type of dependence, and then examine the performance of estimators of the rigidity parameters based on models that ignore its presence. This should be straightforward to implement, however deriving the appropriate likelihood function for the semiparametric estimator in the presence of dependence might be considerably challenging!

Taking a wider perspective, one might consider the use of the framework described here, which allows for the detection of distortions into the shape of the wage growth distributions (WGDs), to investigate features of the wage adjustment process other than downward wage rigidity, that could also produce such distortions that are identifiable. Indeed, we note that this was the case with the investigation for menu cost behaviour in Chapter 2. Another interesting case might be the phenomenon of wage increase diffusion across worker groups, that is, the pressure on wage adjustments to conform with the biggest wage increase. This phenomenon, like downward wage rigidity, produces a shift of probability mass towards higher values of the support of the WGD. In particular, it could be thought of as manifesting itself in a second round of wage adjustments, where it induces pressure on all wage adjustments agreed in a prior round to be revised towards the maximum value agreed then. The challenge that appears to exist in jointly allowing for (the two types of) downward wage rigidity and the phenomenon of wage increase diffusion, is the identification of the separate shifts of mass (i.e. the distortions to the shape of the WGD) that is due to the two phenomena, and, at the same time, of the notional WGD - which, in this case, should be defined as the WGD in the absence of both of these phenomena. This will, certainly, depend on the specific way that is chosen to model the effect of the presence of wage increase diffusion. Furthermore, achieving identification may also require exploiting the presence of heterogeneity in the population.

Bibliography

Abraham, K. G., and J. C. Haltiwanger (1995) ‘Real wages and the business cycle.’ *Journal of Economic Literature* XXXIII, 1215–1264

Akerlof, G. A., and J. L. Yellen (1990) ‘The fair wage-effort hypothesis and unemployment.’ *Quarterly Journal of Economics* 105(2), 255b•“283

Akerlof, G.A., W.T. Dickens, and G.L. Perry (1996) ‘The Macroeconomics of low inflation.’ *Brookings Papers on Economic Activity* 1996(1), 1–59 [60–76]

Altissimo, F., L. Bilke, A. Levin, T. Mathä, and B. Mojon (2006) ‘Sectoral and aggregate inflation dynamics in the euro area.’ *Journal of the European Economic Association* 4, 585–593

Altonji, J.G., and P.J. Devereux (2000) ‘The extent and consequences of downward nominal wage rigidity.’ In *Research in Labor Economics,* ed. S.W. Polachek, vol. 19 (Elsevier Science Inc.) chapter 10, pp. 383–431

Álvarez, L., E. Dhyne, M. Hoeberichts, C. Kwapil, H. Le Bihan, P. Lünnemann, R. Martins, F.and Sabbatini, H. Stahl, P. Vermeulen, and J. Vilmunen (2006) ‘Sticky prices in the euro area: a summary of new micro-evidence.’ *Journal of the European Economic Association* 4, 575–584

Amemiya, Takeshi (1976) ‘The maximum likelihood, the minimum chi-square and the nonlinear weighted least-squares estimator in the general qualitative response model.’ *Journal of the American Statistical Association* 71(354), 347–351

Barwell, R., and M.E. Schweitzer (2007) ‘The incidence of nominal and real wage rigidities in Great Britain: 1978-98.’ *Economic Journal* 117(524), F553–F569

Bauer, T.K., H. Bonin, L.F. Goette, and U. Sunde (2007) ‘Real and nominal wage rigidities and the rate of inflation: Evidence from West German micro data.’ *Economic Journal* 117(524), F508–F529

Behr, A., and U. Pötter (2005) ‘Downward nominal wage rigidity in Europe: A new flexible parametric approach and empirical results.’ CAWM Discussion Paper No. 14

Beissinger, T., and C. Knoppik (2001) ‘Downward nominal rigidity in West German earnings, 1975-95.’ *German Economic Review* 2(4), 385–417

Berkson, Joseph (1955) ‘Maximum likelihood and minimum χ^2 estimates of the logistic function.’ *Journal of the American Statistical Association* 50(269), 130–162

Bewley, T. F. (1999) *Why Wages Do Not Fall During a Recession?* (Harvard University Press)

——— (2004) 'Fairness, reciprocity, and wage rigidity.' IZA DP 1137

Card, D., and D. Hyslop (1997) 'Does inflation grease the wheels of the labor market?' In *Reducing Inflation: Motivation and Strategy,* ed. C. Romer and D. Romer (Chicago: University of Chicago Press) pp. 114–121

Christofides, L.N., and D. Li (2005) 'Nominal and real wage rigidity in a friction model.' *Economics Letters* 87, 235–241

Christofides, L.N., and M.T. Leung (2003) 'Nominal wage rigidity in contract data: A parametric approach.' *Economica* 70(280), 619–638

Christofides, L.N., and P. Nearchou (2007) 'Real and nominal wage rigidities in collective bargaining agreements.' *Labour Economics* 14, 695–715

Christofides, L.N., and T. Stengos (2001) 'A non-parametric test of the symmetry of PSID wage-change distributions.' *Economics Letters* 71, 363–368

——— (2002) 'The symmetry of the wage-change distribution: Survey and contract data.' *Empirical Economics* 4, 705–723

——— (2003) 'Wage rigidity in Canadian collective bargaining agreements.' *Industrial and Labor Relations Review* 56(3), 429–448

Cramton, P., and J. S. Tracy (1992) 'Strikes and holdouts in wage bargaining: Theory and data.' *American Economic Review* 82, 100–121

Crawford, A. (2001) 'Downward nominal-wage rigidity: Micro evidence from Tobit models.' Bank of Canada WP 01-7

Crawford, A., and A. Harrison (1998) 'Testing for downward rigidity in nominal wage rates.' In 'Price Stability, Inflation Targets and Monetary Policy' (Ottawa: Bank of Canada) pp. 179–225

Devicienti, F., A. Maida, and P. Sestito (2007) 'Downward wage rigidity in Italy: Micro-based measures and implications.' *Economic Journal* 117(524), F530–F552

Dhyne, E., L. J. Álvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lünnemann, F. Rumler, and J. Vilmunen (2006) 'Price changes in the Euro Area and the United States: Some facts from individual consumer price data.' *Journal of Economic Perspectives* 20(2), 171–192

Dickens, W., and E. Groshen (2004) 'The International Wage Flexibility Project (IWFP).' (Proceedings of the Final Conference, European Central Bank, Frankfurt Am Main, Germany)

Dickens, William T., Lorenz Goette, Erica L. Groshen, Steinar Holden, Julian Messina, Mark E. Schweitzer, Jarkko Turunen, and Melanie E. Ward (2007) 'How wages change: micro evidence from the International Wage Flexibility Project.' *The Journal of Economic Perspectives* 21(2), 195–214

Dickens, W.T., and L. Goette (2006) 'Estimating wage rigidity for the International Wage Flexibility Project.' Brookings Institution, mimeo

Domencich, T., and D. McFadden (1975) *Urban Travel Demand: A Behavioural Analysis* (Amsterdam: NOrth-Holland)

Du Caju, P. et al (2008) 'Understanding sectoral differences in downward real wage rigidity: workforce composition, competition, technology and institutions.' Paper presented at the EALE 2008 Conference, Amsterdam, Holland

Dunlop, J.T. (1938) 'The movement of real and money wages.' *Economic Journal* 48, 413–34

Fehr, E., and L. Goette (2005) 'Robustness and real consequences of nominal wage rigidity.' *Journal of Monetary Economics* 53(4), 779–804

Fortin, P. (1996) 'The great Canadian slump.' *Canadian Journal of Economics* 29(4), 761–787

Fougère, D., H. Le Bihan, and P. Sevestre (2005) 'Heterogeneity in consumer price stickiness: a microeconometric investigation.' ECB Working Paper, No. 536

Friedman, M. (1968) 'The role of mometary policy.' *American Economic Review* 58(1), 1–17

Friedman, M. (1977) 'Nobel lecture: Inflation and unemployment.' *Journal of Political Economy* 85(3), 451–472

Goette, L. F., U. Sunde, and T. K. Bauer (2007) 'Wage Rigidity: Measurement, Causes and Consequences.' *Economic Journal* 117(524), F499–F507

Goodfriend, M., and R. G. King (1997) 'The Neoclassical synthesis and the role of monetary policy.' Federal Reserve Bank of Richmond Working Paper Series No. 98-05

Groshen, E.L., and M.E. Schweitzer (1996) 'Macro- and microeconomic consequences of wage rigidity.' Federal Reserve Bank of Clevelant WP No. 9607

Holden, S. (1989) 'Wage drift and bargaining: Evidence from Norway.' *Economica* 56(224), 419–432

—— (1994) 'Wage bargaining and nominal rigidities.' *European Economic Review* 38, 1021–1039

—— (1998) 'Wage drift and the relevance of centralised wage setting.' *Scandinavian Journal of Economics* 100, 711–731

—— (2002) 'Downward nominal wage rigidity - contracts or fairness considerations?' Working Paper

—— (2004) 'The costs of price stability: Downward nominal wage rigidity in Europe.' *Economica* 71, 183–208

Holden, S., and F. Wulfsberg (2007) 'Downward nominal wage rigidity in the OECD.' Working Paper

Kahn, S. (1997) 'Evidence of nominal wage stickiness from microdata.' *American Economic Review* 87(5), 993–1008

Kahneman, D., J. Knetsch, and R. Thaler (1986) 'Fairness as a constraint on profit seeking: entitlements in the market.' *American Economic Review* 76, 728–741

Keynes, John Maynard (1936) *The General Theory of Employment, Interest and Money* (London: Macmillan, reprinted 2007)

Knoppik, C. (2007) 'The kernel-location approach: A new non-parametric approach to the analysis of downward nominal wage rigidity in micro data.' *Economics Letters* 97(3), 253–259

Knoppik, C., and T. Beissinger (2003) 'How rigid are nominal wages? Evidence and implications for Germany.' *Scandinavian Journal of Economics* 105(4), 619–641

Kramarz, F. (2001) 'Rigid wages: What have we learnt from microeconometric studies.' In *Advances in Macroeconomic Theory,* ed. J. Dreze (Oxford, UK: Oxford University Press) pp. 194–216.

Layard, R., S. Nickell, and R. Jackman (1991) *Unemployment: Macroeconomic Performancen and the Labour Market,* (Oxford: Oxford University Press)

Lebow, D.E., D.J. Stockton, and W.L. Wascher (1995) 'Inflation, nominal wage rigidity, and the efficiency of labor markets.' Finance and Economics Discussion Series 1995-45. Washington: Board of Governors of the Federal Reserve System

Li, Q., and J. Racine (2007) *Nonparametric Econometrics* (Princeton University Press)

Macleod, W. B., and J. M. Malcomson (1993) 'Investment, holdup, and the form of market contracts.' *American Economic Review* 37, 343–354

Maddala, G.S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics* (Cambridge, UK: Cambridge University Press)

Malcomson, J. M. (1997) 'Contracts, hold-up, and labor market.' *Journal of Economic Literature* 35(4), 1916–1957

McLaughlin, K.J. (1994) 'Rigid wages?' *Journal of Monetary Economics* 34, 383–414

—— (1999) 'Are nominal wage changes skewed away from wage cuts?' *Federal Reserve Bank of St. Louis Review* 81(3), 117–132

Nickell, S., and G. Quintini (2003) 'Nominal wage rigidity and the rate of inflation.' *The Economic Journal* 113, 762–781

Palenzuela, D. Rodriguez, G. Camba-Mendez, and J.A. Garcia (2003) 'Relevant economic issues concerning the optimal rate of inflation.' European Central Bank Working Paper Series, No. 278

Shafir, E., P. Diamond, and A. Tversky (1997) 'Money illusion.' *Quarterly Journal of Economics* 112, 341–374

Shapiro, C., and J. E. Stiglitz (1984) 'Equilibrium unemployment as a worker discipline device.' *American Economic Review* 74(3), 433–444

Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis* (New York, NY: Chapman and Hall)

Smith, J. (2000) 'Nominal wage rigidity in the United Kingdom.' *The Economic Journal* 110, C176–C195

Solon, G., R. Barsky, and J. Parker (1994) 'Measuring the cyclicality of real wages: How important is composition bias.' *Quarterly Journal of Economics* pp. 1–25

Stiglbauer, A. (2002) 'Identification of wage rigidities in microdata - a critical literature review.' In 'Focus on Austria' number 3 (Oesterreichische Nationalbank)

Tarshis, L. (1939) 'Changes in real and money wages.' *Economic Journal* 49, 150–154

Tobin, J. (1972) 'Inflation and unemployment.' *American Economic Review* 62(1), 1–18

Vermeulen, P., M. Dias, M. Dossche, E. Gautier, I. Hernando, R. Sabbatini, and H. Stahl (2007) 'Price setting in the euro area: some stylised facts from individual producer price data and producer surveys.' ECB Working Paper No. 727.

Wasserman, L. (2006) *All of Nonparametric Statistics* (New York, NY: Springer)

Zellner, A., and T.H. Lee (1965) 'Joint estimation of relationships involving discrete random variables.' *Econometrica* 33(2), 382–394