

# Research Statement

Dr Hector Rufrancos

Spring 2023

I am an economist that uses the tools of applied econometrics to study key issues surrounding a number of fields including labour, development economics, political economy and economic history. I am interested in answering questions related to these fields using large or complex datasets ('big data'), often using administrative data, but also combining this with extensive 'shoe-leather' collection methods.

It has become increasingly clear in the last decade that good policy alone does not lead to good outcomes for individuals or society. Instead the credibility revolution and evidence based policy has found itself to be constrained by the messy interplay between politics and institutions. My research agenda seeks to use the tools of the applied econometrician to shine a light on how policy, institutions and individuals interact and how this has implications for the wellbeing of individuals and households.

I describe below my work as organised into two broad headings: 1. Living Standards & Well-being and 2. Institutions & Political Economy 3. Future developments and research plans.

## 1 Living Standards & Well-being

This strand of my research studies how policy affects the living standards and well-being of individuals. This agenda includes work that is a legacy of an ESRC-funded research project in collaboration with Professors Ian Gazeley (LSE) and Andrew Newell (Sussex) and LSE alongside my economic historian colleagues: Rose Holmes, Kevin Reynolds (Brighton) and Cecilia Lanata Briones (Warwick).

This work deals with income inequality in the long-run of the long twentieth century. We have collected datasets which allow us to investigate the dynamics of income inequality and household well-being by exploiting the first household surveys which were conducted. The data span the globe from the 1830s to the 1960s.

In collaboration with my colleagues we collected and digitised some early surveys including the 1953/54 Ministry of Household Expenditure survey. In the 'The Poor and the Poorest, fifty years on: Evidence from British Household Expenditure Surveys of the 1950s and 1960s', It is this kind of 'big old data' which we exploit to describe the wellbeing of households in the period between 1953/54 and 1961. We compare our results with the landmark Abel-Smith and Townsend report studying poverty in Britain in the wake of WW2. We find that contrary to contemporary analysis relative poverty grew, and absolute poverty fell. This was published in JRSSA [ABS 3★]

In 'How hungry were the poor in 1930s Britain?' we digitise and analyse individual household level data from the Ministry of Labour 1937/8 household expenditure enquiry

and those collected by the Rowett Research Institute in 1938/9 for the Carnegie Trust to construct new estimates of food consumption and nutrition re-examining the energy and nutrition available to British working-class households in the late 1930s. We seek to establish the prevalence of malnutrition in the 1930s and its evolution and the nutritional consequences of the school meals and school milk schemes. This was published in the Economic History Review [ABS 4★].

In ‘Escaping from hunger before WW1: the nutritional transition & living standards in Western Europe & USA in the late nineteenth century’ we explore the living standards of American and Western European households in the late nineteenth century. We exploit a high quality microdata to apply a propensity score matching approach to compare similar households. We find that the superiority of American dietary energy over all European countries. British households on average had at least the same energy availability relative to requirements as their American peers, though the gap between American households and other European households was still pronounced. Finally, we provide the first estimates the welfare gain in terms of nutrition to European migrant households to the USA. This was published in Cliometrica [ABS 2★].

In ‘Household survey-based evidence on economic inequality in Britain, 1937-61’ estimate household income inequality for the UK 1937- 1961. We use our household data which shows no movement in inequality below the highest incomes. Secondly, we compare data from 1953/4 to the Family Expenditure Survey for 1961. We find that by 1961, inequality was higher than 1953/4. We find the rise is due to an increase in small, low income, mostly non-working households. This work is currently revise and resubmit in the Economic History Review [ABS 4★]

Along with Gazeley, Newell, Reynolds, Holmes and Lanata-Briones we are working on several other outputs, on the evolution of income inequality for Western Europe 1860-1960, and the evolution of income inequality in Latin America 1914-1968.

In ‘Does the depth of informality influence welfare in urban Sub-Saharan Africa?’ alongside with Egger (UNU-WIDER) and Poggi (AFD), We explore the relationship between household welfare and informality, we propose a novel measure for household informality as the share of members’ activities (hours worked or income) without social insurance. We estimate the cross sectional returns to informality for five African countries. We find that the welfare of mixed formality households is as good as that of fully formal ones. We further corroborate this using longitudinal data where we exploit a novel differences-in-differences approach that carefully (and systematically) selects the appropriate control groups to nullify the selection issue when transitioning household formality portfolios. We find that households do not perceive any welfare differences when transitioning with the notable (and extreme) exception of those switching from fully informal to fully formal. This paper is forthcoming in Oxford Economic Papers [ABS 3★].

## 2 Institutions & Political Economy

This strand of my research studies how politics and institutions affects policy, and in turn how policy affects agents and society.

In ‘The unequal effect of pollution exposure on labour supply across gender’ with Poggi (AFD), Egger (UNU-WIDER), Moro (Stirling) and Schwarz (PIK Potsdam) we exploit high frequency pollution data hourly data on pollution and match this to restricted-access data for

a quarterly labour force survey in Mexico to analyse the daily labour supply of workers in Mexico city over a 5 year period. We exploit some policy thresholds for a pollution alert and employ a regression discontinuity design to document that increases the salience of pollution to estimate the causal effect public policy on labour supply. We then study the contemporaneous effect of pollution exposure at moderate and low levels levels and find that it also reduces labour supply, whilst there are dramatic reductions at high levels of pollution with heterogeneous effects by gender and formality status. This work is ongoing, and is due to be submitted for a grant to update the work through usage of more up to date secure access data. There is a companion piece to this paper that studies the effect of high levels of pollution on crime. This exploits the high frequency pollution data with precisely geolocated administrative crime data for Mexico city. The analysis is carried out a policing quadrant using a high dimensional fixed effects strategy I find that higher levels of PM2.5 are associated with higher levels of violent crime. Yet, when controlling for endogeneity in pollution using the common IV strategy found in the literature exploiting both wind direction and thermal inversions the effect flips. I then show that this is due to the collider bias due to public policy kicking in that shuts down the city at dangerously high levels of pollution.

In ‘Are there gains to be made from joining a union? Evidence from Mexico’, I seek to quantify the union wage premium and loss associated with transitioning to/from a union. The labour economics literature has highlighted the tricky issue of estimating the average treatment effect due to selection effects among union members. I recover the ATT implicit in joining / leaving a union by carefully selecting treatment and control groups. I find that the ATT is overestimated between the typical cross section estimates by 50%. In addition I estimate the non-wage benefit losses and gains associated with union switches. This was published in BJIR [ABS 4★]

In ‘The impact of University reopenings on COVID-19 cases in Scotland’ alongside colleagues Moro (Stirling) and Moore we analyse the unintended consequences of allowing student in-migration into student halls in Autumn 2020. This piece of work was highly policy salient and was written with the express goal of leading to evidence-based policy for Autumn 2021 university in-person reopening. To do so we hand collected through laborious ‘shoe-leather’ research the location of all student halls of residence in Scotland both university owned and those private provided. We exploit these data to identify student neighbourhoods and match these data with high frequency daily data on COVID-19 cases. We use this to estimate the effect of the start of semester on daily COVID-19 cases. We sidestep the concerns raised in the novel literature on Differences-in-differences by focusing on valid comparison groups and estimate a local differences-in-differences focusing on comparing student neighbourhoods with their 1km nearest neighbouring neighbourhoods. We find that for the cohort of universities whose term began on 14/09 the impact of students moving in peaked at 400 cases per 100k of the population 1 week after the start of term. We further investigate if there are spillover effects from students moving in on the nearby locations and find a persistent increase in the nearby localities.

In ‘Bismarck in the kitchen? The effect of early social protection on nutritional intake’ Gazeley (LSE), Newell (Sussex), Cowan, Duncan and I use high-quality historic microdata for 1890 and 1907/08 combined with modern identification methods to analyse the nutritional impact of the rollout of the first pensions system world-wide. We employ a series of Regression Discontinuity Designs along the threshold for pension eligibility to show that households foodshare is unaffected by pension eligibility, yet energy availability for the poorest households is found to be same as for those households who are 10% wealthier. We further exploit our

historic microdata and employ a differences-in-discontinuities design to further underscore that our findings are due to pensions eligibility and the subsequent behavioural response by households.

In ‘Institutions in the fast lane? Strike Petitions and the Electoral cycle in Mexico’ with Sas (Stirling) we investigate the preconditions under which institutional change actually leads to changes in which players play the game. In this paper we investigate the behaviour of trades unions in the post-democratisation period in Mexico. Using a sharp regression discontinuity design we find that narrowly-won mayoral elections by right-(and left-)wing candidates cause increases in strike threats two years after an election, during the campaign for the following election. We develop a political economy model of union influence in electoral competition and find that when electoral institutions become more democratic, and political parties can count to a lesser extent on a guaranteed support base, the importance of relying in the unions to win elections will increase. We exploit unique access to ‘big administrative data’ covering the universe of all strike threats in Mexico from 1991–2012. This research is important for three reasons. Firstly, it is the only paper which explicitly looks at strike threats, the bargaining signal, rather than strikes, the failed outcome of bargaining. Second, it contributes to the scant literature on the role of unions and politics, by providing the first causal evidence of the role of the political cycle on private sector unions. And finally, this paper also contributes to the growing literature on the persistence of poor institutions by providing a plausible channel by which vested interests ensure that even though the ‘rules of the game’ have changed, the behaviour of ‘players’ remains unchanged.

### 3 Future Research Plans

At present I am Developing a large grant ‘What Drives Food Insecurity in the UK? Austerity, Benefits & Precarious Work’, if successful this will be a three year project with at least three main outputs. This research has a high potential for impact as it dovetails with the present cost-of-living crisis. The co-Is are Moro (Stirling), Stowasser (Stirling). I have put together the team, and have secured the institutional collaboration of The Trussell Trust who will give us access to ‘big administrative data’ on the universe of all food bank users in the Trussell Trust Network food banks for the period 2010–Present. The data collaboration will go ahead regardless of the funding situation. This project has two aims. First, we will investigate the causes of food insecurity and food bank usage in the UK. Specifically, we will study the extent to which several government policies are responsible for the growth in food insecurity. Second, we will communicate our findings to relevant stakeholders in academia, the charity sector, and the policy sector with the goal to better inform political discourse and ultimately improve policymaking. To that end, this project seeks to answer three research questions: (1) How have local-authority austerity measures affected food bank usage? Previous research on the effect of austerity has focused on national policy. However, differences in how local authorities implemented austerity policies provide “natural experiments” that we can use to identify the casual effects of austerity. (2) How has the growth of zero-hour contracts, and work insecurity more broadly, affected food bank usage? Given that the prevalence of zero-hours contracts is likely highly underestimated (Farina et al., 2020), there is a real risk that this potentially important driver of food bank demand is unjustly ignored by policymakers. (3) How have changes to the welfare system, in particular the introduction of Universal Credit and its benefit-sanction regime, affected food bank usage? This part of the analysis will

exploit several quasi-exogenous shocks to the welfare system (such as the staggered rollout of Universal Credit and its temporary uplift during the Covid-19 pandemic) to disentangle transitory and persistent effects of welfare reform. Food insecurity is a topic of contentious public debate. The COVID-19 pandemic and the current cost-of-living crisis have further exacerbated nutritional inequality in the UK. These circumstances mean that reliable, causal evidence that identifies the drivers of food bank demand has never been more crucial and will be essential in supporting the development of effective policy responses aimed at alleviating poverty.

Alongside my colleague Schroeder (Birmingham) we working to evaluate the impact of the roll-out of the drinking bye-laws in Glasgow. We will use ‘shoe-leather methods’ to painstakingly digitise court records to generate a historical spatial geolocated dataset of crime records for 1990s Glasgow. We will exploit this data to quantify the decrease in violent crime due to the drinking bye-laws.

Exploiting this same approach to data gathering I have a project sketched out seeking to digitise trial records in turn of the century Glasgow to determine if there was judicial bias in sentencing along the sectarian dimension. There is a growing literature in Economics (e.g., see Anwar et al., 2018; Ash et al., 2022) demonstrating that in-group dynamics may spillover into the functioning of institutions such as the judiciary. This work would have two outputs namely the paper, but also the research data. I expect this to be of use for future research.

Another historical institution / policy paper looks at the role of Carnegie libraries in the UK on the hollowing out of communities and social mobility & innovation. This work aims to exploit the ESRC-funded cross census linkage for the microdata for the 1890, 1900 and 1910 census. I aim to look at the roll out of the Carnegie libraries in the UK. Early work (on aggregated census data) demonstrates that occupational structure of cities that got Carnegie Libraries changes, and in some instances, it is quite successful in promoting social mobility, leading to the unintended consequences of brain-draining towns, whilst central locations such as London benefit from this move. This work would seek to rigorously apply current causal identification methods to big old data to show the impact and importance of public libraries. This work in the next five years would be submitted to the ESRC for a secondary data analysis initiative grant.

More broadly, in the next five years, I am interested in further pursuing the broad research agenda outlined above spanning multiple fields, but always ensconced in careful causal identification mixed with big data and bespoke data collection. I look forward to the opportunities which open themselves to me and finding more collaborators to further enhance these agendas.