RegData: Australia

Patrick A. McLaughlin, Jason Potts,
and Oliver Sherouse

MERCATUS WORKING PAPER

3434 Washington Blvd., 4th Floor, Arlington, Virginia 22201
www.mercatus.org

Patrick A. McLaughlin, Jason Potts, and Oliver Sherouse. "RegData: Australia." Mercatus Working Paper, Mercatus Center at George Mason University, Arlington, VA, June 2019.

Abstract

In this paper we introduce RegData Australia (RDAU1.0) and present some preliminary and comparative findings using this new panel. RDAU1.0 applies the RegData method to create a unique Australian database that extends from 1997 to 2012. RegData uses text analysis to quantify restrictive clauses in legislation, significantly improving the accuracy of measurements of regulatory incidence. RDAU1.0 extends and adapts the RegData methodology to Australian regulations and legislation. We use RDAU1.0 to capture broad patterns in Australian regulation, and we compare these data to RegData findings from other regulatory jurisdictions, including the federal government in the United States and several US state governments. A preliminary analysis yields relational evidence consistent with previous researchers' hypothesis that the extent of regulation will be determined by the size of the market because of the fixed costs of regulatory production. This hypothesis suggests that regulatory volume in a specific jurisdiction will scale as a function of the jurisdiction's population. We examine RegData metrics of regulation for 23 different jurisdictions, including the federal governments of Australia and the United States and the state governments of 21 American states, and find a positive and significant correlation between regulatory volume and population.

JEL codes: K2, L5, N4, Y1

Keywords: regulation, RegData, Australia

Author Affiliation and Contact Information

Patrick A. McLaughlin
Director of Policy Analytics
Senior Research Fellow
Mercatus Center at George Mason University
pmclaughlin@mercatus.gmu.edu

Jason Potts
Professor of Economics
Royal Melbourne Institute of Technology
jason.potts@rmit.edu.au

Oliver Sherouse
Former Policy Analytics Lead
Mercatus Center at George Mason University
oliver@oliversherouse.com

Acknowledgments

The authors thank Darcy Allen, Aaron Lane, and Jonathan Nelson for their excellent research assistance on this project.

This paper can be accessed at http://www.mercatus.org/publications/regulation/regdata-australia

RegData: Australia

Patrick A. McLaughlin, Jason Potts, and Oliver Sherouse

1. Introduction

There are few parts of a modern market economy that are not regulated in some way. Economic theory has multiple explanations for why this occurs; they are based on who benefits from regulation. The public interest model predicts that consumers benefit when governments regulate to correct market failures (Pigou 1932). The private interest model predicts that businesses benefit when government regulation limits entry into markets (Stigler 1971, Djankov et al. 2002). Economic theory also seeks to explain why governments prefer regulation to other forms of social control of business, such as control through public ownership or through courts (Glaeser and Shliefer 2003, Mulligan and Shleifer 2005, Shleifer 2010). Economic theory also endeavors to explain the costs of regulation, much of which are hard to see because of how regulation restricts innovation (Thierer 2014).

However, scientific analysis of the economic effects of regulation remains challenging because of the difficulty inherent in measuring what is naturally a qualitative object. Regulations are rules, prescriptions, or orders that are issued by governments following their democratic whims. Regulations are written by lawyers using legal language and enacted by agencies with their own agendas, and they carry the force of law. Their effects are easy to theorize about but hard to measure.

Of course regulation can be quantified—that is, as the number of regulations issued—and this measurement can be elaborated by counting the number of words or pages in the regulation (Coglianese 2002, Dawson and Seater 2013) or examining the file size of the

digitized version of the regulatory code (Mulligan and Shleifer 2005). But there is a very imperfect mapping between the number and size of regulations and their economic consequences. Researchers therefore often prefer to focus on single regulations or regulatory targets, and to base analysis on a quantified event, such as market entry or job separation, tied to the introduction of or changes to a particular regulation or set of regulations (e.g., Greenstone 2002). Related approaches use natural experiments to the same effect (e.g., List et al. 2003). Another strategy is to measure regulation indirectly, such as through estimates of firm- or industry-level compliance costs (e.g., Jaffe et al. 1995).

These methods are adequate for some questions, but many lines of inquiry remain closed. In particular, it is very difficult to study the effect of single regulations on multiple industries, or of multiple regulations on single industries. Comparative regulatory analysis—analysis of the effect of regulation between industries or across countries—remains a mostly speculative domain. Studies of the demand for and supply of regulation are largely constrained to analysis of single regulations, and thereby ignore cumulative effects (McLaughlin and Greene 2014).

Recently, however, researchers at the Mercatus Center at George Mason University have made a significant advance in solving this underlying measurement problem with the development of a numerical database panel and an associated method, called RegData (Al-Ubaydli and McLaughlin 2014, McLaughlin and Sherouse 2016). RegData is a new approach, intended as an open access research protocol, that uses machine-learning-based textual analysis of the *Code of Federal Regulations* (CFR) to measure the actual restrictive content of legislation and regulations as indicated by words such as "shall," "must," "should," "prohibited," and so on (rather than measuring the container of regulation, for instance measuring page counts or file sizes). Using industry-specific training documents, RegData goes further, identifying the

relevance of regulatory text to different industries as defined by the two- through six-digit levels of the North American Industry Classification System (NAICS). Al-Ubaydli and McLaughlin (2014) explain that "RegData is the first panel of federal regulation for the United States annually for the years 1997–2012 that permits within-industry and between-industry econometric analyses of the causes and effects of federal regulations."

The purpose of this paper is to introduce an extension of RegData to create a further panel by applying the method to a new country: Australia. We will explain the problems and challenges this involved and outline the initial findings. The purpose of creating this extension is twofold. First, this new panel furnishes a new dataset that can be used in the study of Australian regulation and its effect on the Australian economy, with decomposition to the level of different acts of Parliament and regulatory agencies.[1] This sort of analysis has not previously been possible, because until now such a panel has never existed. Obviously, this dataset is principally of interest to Australian scholars.

However, there is a second, broader purpose in that we intend that Australia will be the first of many subsequent RegData panels. Such development will enable cross-country analysis, making possible a new field of comparative institutional and regulatory economics. This multi-country RegData database will likely be instrumental in scholars' efforts to analyze the regulatory effect on global patterns of trade and production and on the location of businesses and the global distribution of entrepreneurial action, and therefore of financial capital. This sort of analysis will also contribute to empirical work in the new comparative historical political economy (Boettke et al. 2005, Boettke et al. 2013, Hodgson 2015), comparative law and economics, comparative

[1] We intend that subsequent work will seek to decompose these findings by industry using Australian and New Zealand Standard Industrial Classification (ANZSIC) codes (which are similar, but not identical, to NAICS codes). This will require the generation of a new set of training documents for the RegData program based on documents with known direct and strong relevance to Australian industries at two- and four-digit ANZSIC classifications.

international law, and new development economics. In this paper we have endeavored to set out the main issues and problems that accrue in adapting and applying the RegData method to a novel political and institutional context, and some of the resolutions we achieved.

The first version of RegData was released in 2012.[2] Subsequent versions included RegData 2.0, described in Al-Ubaydli and McLaughlin (2017), and RegData 2.2, described in McLaughlin and Sherouse (2019). The primary difference between 2.2 and earlier versions is the addition of machine-learning algorithms to improve the accuracy of industry classification (as described in McLaughlin and Sherouse 2016). This paper will not provide a detailed overview of RegData, for which the reader is referred to Al-Ubaydli and McLaughlin (2014), the appendices in Al-Ubaydli and McLaughlin (2014), and McLaughlin and Sherouse (2016). While the method used in RegData 2.2 is the same as the method we employ in the Australian context in this paper, the political, regulatory, and economic context of Australia differs sufficiently from that of the United States that the process of creating a new Australian panel is far from straightforward.

Section 2 reviews that Australian context and provides an overview of the adaptation of the RegData 2.2 method and panel to that new context. We call this version of RegData RDAU1.0, because it is the first version of RegData in Australia. Section 2 describes the creation of RDAU1.0 in such a way as to guide researchers seeking to use RDAU1.0 for analysis of Australian regulations and their economic effect, and also to guide researchers from other countries seeking to implement RegData to create a new local panel. Section 3 presents some

[2] As Al-Ubaydli and McLaughlin (2014: footnote 4) explain, "RegData was first introduced in a working paper published in July 2012; see http://ssrn.com/abstract=2099814. The version of RegData we introduce in this paper contains several improvements over the July 2012 version. Improvements include data for years 2011 and 2012; data for NAICS four-digit industries; search-term weightings derived from Google's Ngram database; scalable granularity for CFR search results, ranging from CFR title-level results to CFR paragraph-level results; and regulatory agency and sub-agency-specific search results."

preliminary overview findings that indicate the type of data and results that RDAU1.0 shows. Further discussion and conclusions follow in section 4.

2. RegData in Australia—RDAU1.0

2.1. The (Lack of) Measurement of Regulation in Australian Policymaking

The Australian economy has long been regarded as one of the world's best-performing economies.[3] Part of this success is attributed to responsible public fiscal and monetary management, which has resulted in relatively low levels of inflation and public debt. Institutional quality, including those institutions that are included in measures of economic freedom such as property rights, judicial independence, and sound money, is another widely acknowledged factor (Gwartney et al. 2016). But the other major part of this conventional story has to do with the immediate and lasting effects of economic deregulation.

Australia was one of several countries around the world—along with the United Kingdom under the Thatcher government, the United States under the Reagan government, and New Zealand under the Lange government—that, through economic reforms that began in the late 1970s and early 1980s, attempted significant reversal of the growth of the planned economy and the regulatory state. Each country undertook this process in a different way. In Australia, after the enormous growth in the size and reach of government that followed the election of the Whitlam government (1972–1975) and remained largely in place during the Fraser government (1975–1983), serious economic reform began with the Hawke government (1983–1991). The conventional story regarding the reforms begun under the Hawke government is that they deregulated much of the Australian economy—floating the dollar, privatizing public-sector

[3] For a long-run historical perspective, see Butlin et al. (1982).

industries, deregulating finance, cutting tariffs and taxes. This agenda more or less continued through the Keating era (1991–1996) and then the Howard era (1996–2007).

During the Rudd-Gillard era (2007–2014), a broad slowdown in the Australian economy has been blamed on a number of external factors, such as the global financial crisis and reduced exports to China, as well as on internal factors, such as growing wages and other costs of business in Australia. Productivity growth in Australia has also weakened, a development in part attributed to a rise in economic regulation (Banks 2004, Davis and Rahman 2006, Berg 2008). This was particularly associated with ever more restrictions in labor markets, transport (Everett 2006), and the building and construction sector (Berg et al. 2005), and with the chilling effect of increasing environmental legislation on investment in mining and agriculture (Davidson and Elliston 2005, Banks 2005).

The arguments that led to the various waves of deregulation in Australia were largely the outcome of a war of ideas shaped by elements of economic theory, political rhetoric, and practical experience (Fels 2004, Moran 2006). And the subsequent reregulation of Australian markets and industries over the past decade or so has occurred mostly owing to a realignment of political forces that were focused tightly on concentrated benefits of regulation and were able to ignore broader and unseen costs. It is reasonable to say that at no point were the decisions leading to these policy changes (decisions by voters, politicians, regulators, and other stakeholders) based on theories derived from empirical evidence about the aggregate and differential consequences of regulation, and of regulatory changes. For that matter, even the conventional wisdom about historical changes to regulation—for example, that the Hawke government undertook major deregulatory efforts—mostly comes from anecdotal evidence, and

comprehensive measurement of regulation may show this characterization to be incorrect. The reason, of course, is the paucity of rigorous panel data of regulation in Australia.

The upshot is that the practice of regulation in the Australian economy, both by those broadly in favor of it and by those against it, has been at best only weakly informed by rigorous empirical evidence. There have been high-level acknowledgements of the cost of the regulatory burden in Australia, but only very approximate measurements (Berg et al. 2005: 15–17). While rigorous economic theory and evidence should not be the only consideration guiding political actions, it seems plausible that Australian policymakers would benefit from better data on regulation. Moreover, comprehensive and objective metrics of regulation in Australia would offer to researchers the ability to delve into the causes and effects of various regulatory trends and policies. We hope that the new data and tools introduced in this paper will make it possible to conduct better empirical analysis of regulation and its effects, which in turn will have the potential to guide better decisions about regulation in Australia.

2.2. Adapting RegData 2.2 to the Australian Legislative and Regulatory Context

The Australian political, legislative, and regulatory system differs in both obvious and subtle ways from the US system.[4] These differences affected how RegData 2.2 was able to be implemented as RDAU1.0.

The first important difference concerns the way in which regulations are made and enacted. Australian regulations originate in parliamentary acts, and then through the agency of a federal minister, rather than through the delegation of rulemaking authority from Congress to

[4] At a general level, Australia, like the United States, is a bicameral federal democracy, but Australia is a constitutional monarchy rather than a constitutional republic. The governor-general (who represents the monarch) is the Australian head of state, and the prime minister is the leader of the elected government. Australians vote for a political party, not for a president. Parliament in Australia parallels Congress in the United States.

various regulatory agencies, as in the United States. This means that, rather than attributing regulations to a particular agency, as RegData 2.2 does in the United States, RDAU1.0 attributes them to particular acts.

There are two categories of Australian legislation: acts and legislative instruments. Acts of Parliament are the primary category, made by the Australian Parliament. According to the Constitution of Australia, each draft act (a "bill") must pass through both houses of the Parliament and receive royal assent before it can become law. Once an act is in place, it can generally only be amended or repealed by another act. Legislative instruments are laws on matters of detail made by a person or body authorized to do so by an act of the Parliament. They may be called regulations, rules, and many other names, but their label is not important. What matters is that they have force of law. For example, section 504 of the *Migration Act 1958* authorizes the governor-general (in practice, the responsible minister) to make regulations about particular matters. These regulations are contained in the *Migration Regulations 1994*. From a political and legal perspective, the act and the regulations exist as one regulatory regime, as a set of programmatic rules that work together to regulate a targeted domain.

A second important difference between the Australian and US systems affects the form of the data inputted into the RegData programs. In the United States, there are two highly useful publications that summarize and explain federal regulations and regulatory activity. First, there is the *Federal Register*, which is the US government's journal of daily bureaucratic activity, including proposed and final regulations. Page count measures, such as Coffey et al. (2012), often use this instrument. Second, there is the *Code of Federal Regulations*, which is published annually. RegData uses the CFR as input for text analysis because it contains only the stock of

final federal regulations, without the extraneous explanatory preambles and proposed regulations contained in the *Federal Register*.

Unfortunately, no such summary bureaucratic administrative documentation exists in Australia. In consequence, the data fed into RDAU1.0 have been manually collected from the official online repository of legislation of the Commonwealth of Australia, *ComLaw*.[5] This online database contains current legislation and a limited amount of historical legislation. There was no practical method available to access the entire corpus of Commonwealth law in a way that would approximate the input of the CFR.[6] To implement RDAU1.0, we therefore manually downloaded the entire corpus of Commonwealth legislation and regulation—1,800 acts and 870 regulations—from the *ComLaw* website.[7]

There are several consequences of these differences. One is that, because successive Australian parliamentary acts update each other, with the superseded version becoming a historical act, there is a risk of double- or triple-counting their content (e.g., words or regulatory restrictions). Furthermore, the majority of historical acts do not have listed end dates, which complicates endeavors to track the origin of a regulation. The reason this information is important is that sometimes, usually when governments change between versions of an act, the name of the act changes. Initially, it seemed that the RegData system dealt well with updated

[5] See https://www.comlaw.gov.au/. Data were downloaded between March 13 and 15, 2015.
[6] This point was elucidated by the Office of Parliamentary Counsel, which confirmed that each act and regulation would need to be downloaded and saved individually. The manual downloading and inputting of *ComLaw* files was a significant but unavoidable initial investment in creating RDAU1.0. We have made some progress in automating this, which will make this process more efficient for subsequent countries.
[7] Two initial technical aspects are important to note about these data. First, while the data were collected from the official government website, there were some obviously obsolete documents. After the initial download took place, we manually sifted through the data to remove these wherever possible. Second, some Australian federal parliamentary documents that are too large are split into several volumes (e.g., the *Corporations Act 2001* has five volumes). We reaggregated the volumes of each act and regulation into a single analytical unit (e.g., the *Corporations Act 2001* is aggregated into one document containing all five volumes) to ensure consistency of treatment.

versions of the same act. However, without the end-date entry, the system appeared to have no way of telling whether an act was still current or whether it had been replaced by an act with another name. For example, in the initial analysis, the *Industrial Relations Act 1988*, the *Workplace Relations Act 1996*, and the *Fair Work Act 2009* all showed up as being current in 2014—which is incorrect, because each successive act replaced the former. That is, the *Workplace Relations Act* replaced the *Industrial Relations Act*, and the *Fair Work Act* replaced the *Workplace Relations Act*. This inability to distinguish between current and obsolete legislation was initially problematic because of changes in the *ComLaw* codes. While the problem only affects a handful of acts, these acts are often highly restrictive pieces of legislation. Enabling RDAU1.0 to recognize start and end dates of acts and regulations avoided this double-counting issue, and furthermore allowed a long period of historical comparison to emerge.

Another consequence of the differences between Australia and the United States concerns the effect of Australia's federal and state jurisdictions. Australia is a federation of six states and two self-governing territories, based on the Westminster system. The Australian Constitution divides powers between the Commonwealth and the states. A list of Commonwealth powers, although not complete, is found in section 51 of the Constitution. Generally speaking, the residue of powers not vested in the federal Parliament remains with the states.[8] Major portions of state legislation concern areas such as service delivery (e.g., healthcare, education, public transport), law and order (e.g., crime, antidiscrimination policies, policing, corrections, courts and tribunals), and infrastructure. Other areas of state legislative activity include occupational health and safety legislation, local governments, urban planning, and environmental approvals. The data collected for RDAU1.0 do not capture acts and regulations passed by state and territory

[8] Australia's two territories—the Northern Territory and the Australian Capital Territory (which is the location of Parliament)—are both administered federally.

parliaments. This is a particular limitation of the initial version of RDAU1.0, but provides a rich scope for further research as this research program develops. The immediate implication is that this limitation to federal regulations means that analyses based on RDAU1.0 underestimate the effect of regulation in Australia.

A further issue with the historical acts and regulations is the availability of the data online. Discussions with the Office of Parliamentary Counsel revealed that acts and regulations were progressively uploaded to the online repository in the late 1990s and the first few years after 2000—and then updated as relevant. The result is that data for repealed acts and ceased regulations between 1901 and the 1970s do not appear to be present in the online repository. This makes sense, because these data would need to be scanned manually from printed materials, and there would not have been a use for that until this project.

Another issue we encountered is that some larger acts are broken up into volumes. For example, the *Corporations Act 2001* comprises of 2,860 pages, spanning five volumes. Initially, these volumes showed up as separate acts—causing us to underestimate the effect of this legislation. To address this problem, we manually identified the volumes of these large acts.

2.3. Corpus Quality Checks

In order to ensure the integrity of our datasets, we put our produced data through a series of quality checks. Once the metadata CSV file is created, most of these checks can be performed by analyzing the metadata file rather than analyzing the full text corpus.

One of our first checks is to determine the minimum word count of the corpus and how many documents have that word count. If too many documents have a low word count or zero words, this suggests that the underlying text quality may be compromised. This problem is often caused by difficulties converting the original documents into digitized text.

Another check we perform is to examine the index labels of the documents and look for red-flag words that may indicate that certain documents should not be included in the corpus. For example, if the name of the document includes words like “repealed,” “cancelled,” or “expired,” the document should not be included in a corpus intended to represent the most up-to-date version of the regulatory code. Often these documents have a low word count because they only include a sentence or two explaining that the rule has expired or been repealed. These kinds of documents should be filtered out of the corpus during either the download portion or the cleaning portion of the dataset production.

Another important quality check we perform is to check for duplicates in the metadata file. Duplicates are often indicative of one of several potential problems. First, the index may not be specific enough to capture differences between documents. For example, regulatory codes may have multiple appendices in the same section of the code. If the index is not specific enough to capture the difference between these appendices, they will appear as duplicates in the metadata file. Second, duplicates may be caused by different versions of the regulatory code mixed together in the same source. Third, duplicates may indicate human error during the processing of the text. Out-of-date text may not have been deleted or overwritten when new text was created.

Finally, if the elements of the corpus index are intended to be numeric (i.e., integers or floats instead of text), we can programmatically force the values to be numeric. Errors that may occur suggest that the index has text or other unwanted characters present. This problem should be fixed by organizing or cleaning the index values in a way that allows them to all be numeric.

3. Comparative RegData Findings

Until now, time series estimates of Australian regulation were limited to page number counts (e.g., Berg et al. 2005, Berg 2008, Novak 2013). The preliminary results from RDAU1.0 now enable, by capturing restrictiveness, a new time series portrait that is a more detailed measure of the real state of regulation in Australia. Furthermore, the data can be decomposed into origin of regulation by act. On its own, this provides new empirical measures of the dynamics of regulation in Australia. But because RDAU1.0 uses the same methodology as RegData 2.2, it also enables comparisons with findings from other jurisdictions where RegData has been applied, such as the United States and Canada. We draw out some preliminary inferences that contribute to a new comparative regulatory economics.

The RegData method of textual analysis counts the number of keywords likely to impose binding constraints on citizens or businesses. There is a simple underlying principle: regulations influence society by modifying the choice sets of individuals. By accounting for the actual measure of restrictions, we can quantify the source and strength of the effect of a regulation. RegData is a custom-built computer program developed to count a carefully selected suite of verbs and adjectives. Following Al-Ubaydli and McLaughlin (2014), RDAU1.0 uses five such words: "shall," "must," "may not," "prohibited," and "required." These are the "regulatory restrictions" that measure the extent to which a regulation prohibits or requires certain acts or behaviors by individual Australians or businesses. The outputs of the program are therefore "restriction counts" within a given document of the federal corpus.[9]

[9] Interestingly—and to our knowledge, completely independently of the RegData project—at least two governments have developed and used metrics of regulation that are remarkably similar to the regulatory restrictions metric used in RegData. British Columbia began inventorying "regulatory requirements" in 2001, and the Canadian province has since used this metric in a successful red-tape-reduction and regulatory budgeting program (see Jones 2015 for details). The state of Queensland, Australia, also discussed a metric of "restrictiveness" as a way of measuring the volume of regulation and as a proxy for total burden. Queensland's approach was explicitly modeled after the British Columbia approach (Queensland Competition Authority 2012).

A comprehensive analysis of Australian regulation comparable to the current US version of RegData (i.e., a RDAU2.0) would further include decomposition of regulatory incidence by industry, using codes developed for the Australian and New Zealand Standard Industrial Classification (ANZSIC) at the two- and four-digit levels.[10] An initial version of these estimates has been created, but it is limited by two factors and further research will be required to enable trust in the method and comparison. First, while ANZSIC and NAICS are similar, they are not identical. Second, the RegData method uses training documents sourced from material that is tightly correlated with each industry at the NAICS two- through six-digit levels (e.g., industry reports). Our initial tests of Australian industries used the North American training documents, so the results, while plausible, are accurate to an entirely unknown degree. Hence we have not included these results here, and future research will seek to explain and explore industry-specific results, both as a time-series analysis for Australia and as a US-Australian comparison.

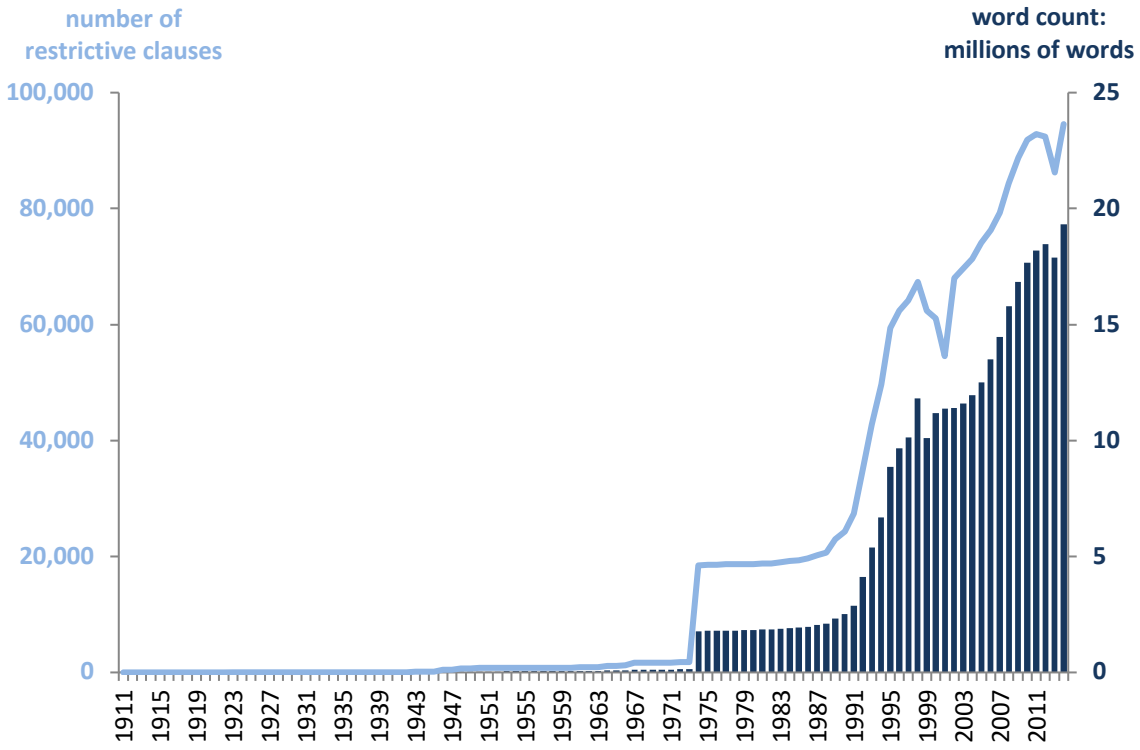
3.1. Comparative Regulation through Time

Figure 1 presents the main summary findings of restrictive clauses in Australian federal legislation through the period 1911–2014. The word count and restrictive clauses data in figure 1 extend the extant page count data (Berg 2008, also see Novak 2013: 28) on aggregate time series Australian regulation measures. Page count data, while the easiest to assemble, obviously provide the least accurate measurement of the real restrictive content of legislation and regulations, because page counts are affected by page sizes, fonts and font sizes, the extent of

[10] All RegData datasets—including RegData Australia as well as the companion RegData datasets for the United States and Canada—are comprehensive in the sense that they include data for all actual regulations (i.e., administrative law). However, RegData does not include any information about guidance documents or other regulation-related documents and actions. It is certainly plausible that some de facto regulation occurs in the form of guidance documents, memos of interpretation, or other forms of "soft law." While the quantification of soft law is beyond the scope of the RegData project, we hope that future research can address it.

tables, and other peculiarities of layout. Word count data are better, because word counts are closer to a measure of legislative and regulatory content. However, word counts also include explanatory content, preambles, appendices, and other extraneous material. Word counts also don't account for differences in style, or—more importantly—scope. The closest measure to real regulatory effect is restrictive clauses, but it is also the most computationally intensive (i.e., the highest cost) to measure.

Figure 1. Restrictive Clauses and Word Counts in Australian Federal Legislation, 1911–2014



Source: Produced by the authors using the RDAU1.0 dataset.

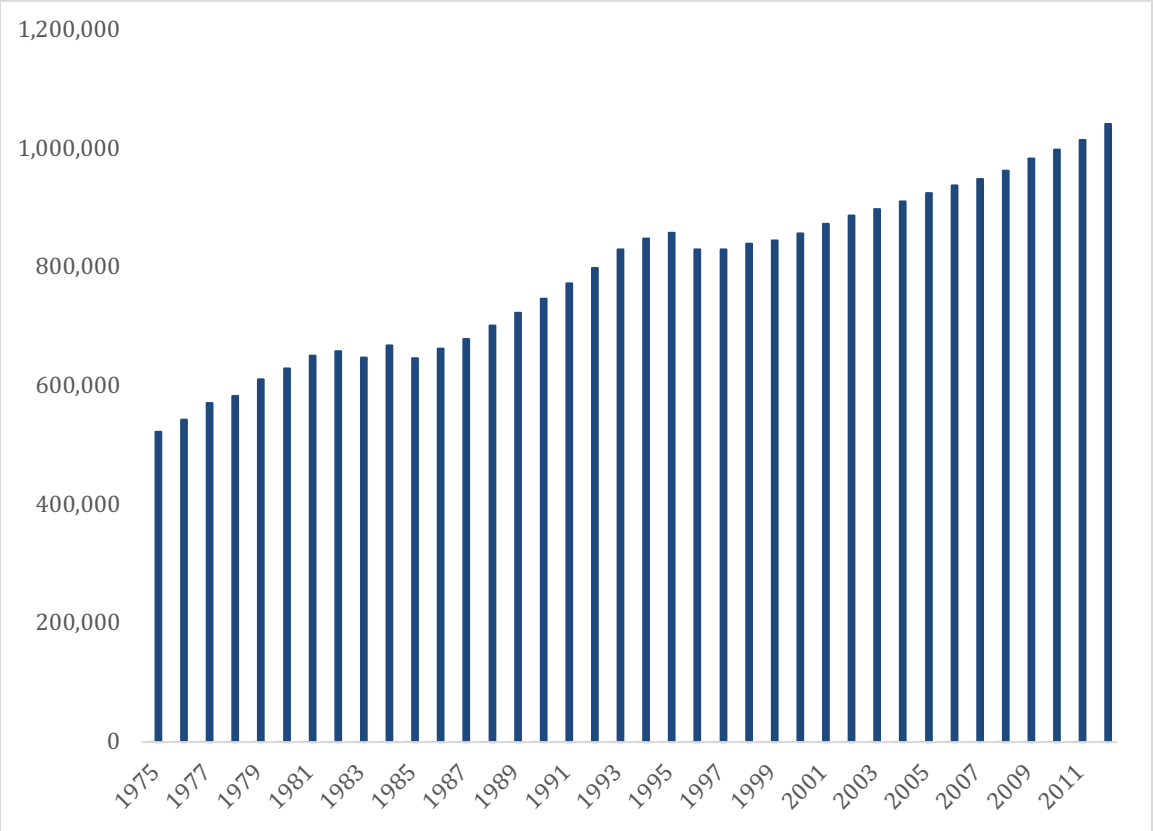
It is notable that page count, word count, and restrictive clauses all more or less move together. So each validates the other. But they are not perfectly correlated, and the presumption

should be that the restrictive clauses measure is the closest to the true measure, followed by word count, then page count.

On the basis of restrictive clauses, we can immediately observe periods of relative stability in Australian regulatory growth (the 1960s, the 1980s, 2009–2012), followed by periods of rapid growth (1972–1975, 1991–1996, the decade following 2000). There are also seemingly interphase periods (1996–2001) that fluctuate. The page count and word count data point to these patterns but do not show them clearly. The restrictive clause data seem to indicate the existence of such distinct dynamic phases. It is unclear at this stage whether these are correlated with political or business cycles, or with demographic or other exogenous causes.

How does the Australian RegData compare with the US RegData? Figure 2 shows the restrictive clauses measure over the period 1975–2012.

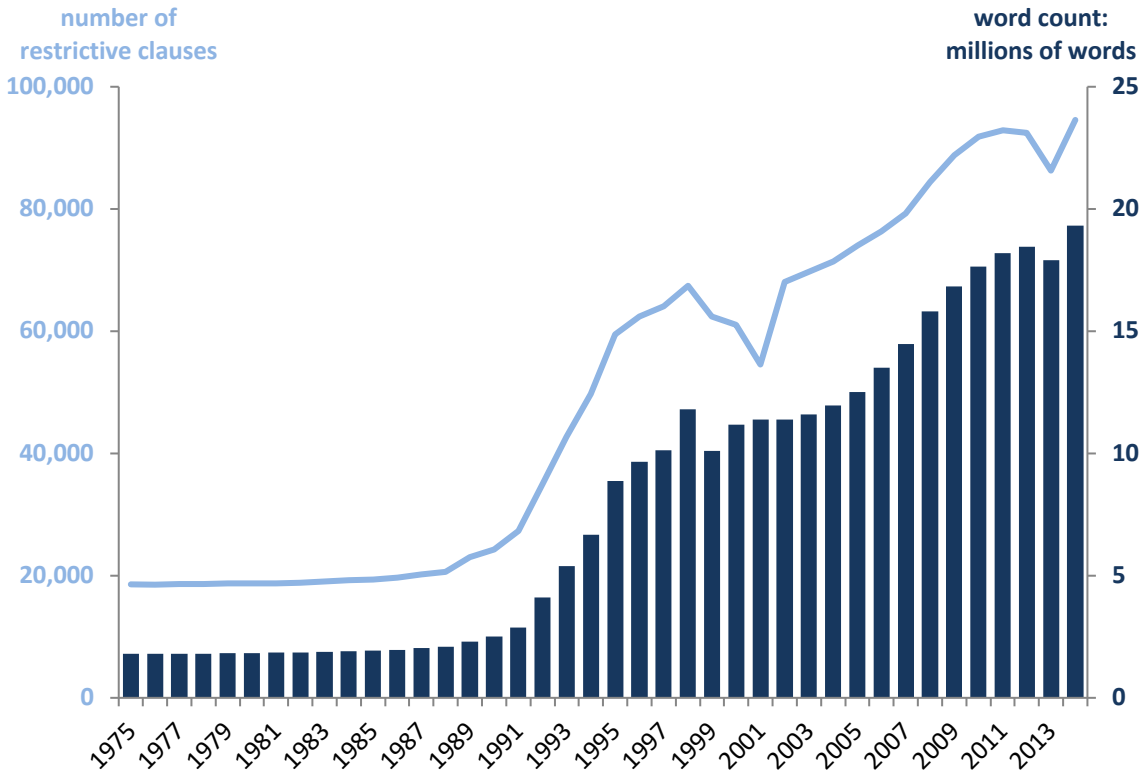
Figure 2. Cumulative Number of Restrictions in US Federal Regulations, 1975–2012



Source: Produced by the authors using the RegData 2.2 dataset, available at QuantGov, https://quantgov.org/regdata-us/.

While the Australian RegData series runs from 1911, we are most confident of the quality of the data from the late 1990s onward because regulatory data were reliably digitized only from this point onward, a circumstance openly conceded by the Australian parliamentary office. The parliamentary office observes that there is little demand for digital regulatory data prior to the late 1990s, and so it has had little incentive to archive them in digital form. We expect that this will be a problem in many other countries as well. So a strict comparison between Australian and US data is of uncertain quality from 1975 to the late 1990s and is of high quality only from 1997 onward. By truncating figure 1 (eliminating data before 1975), we arrive at figure 3, which is comparable with the American data shown in figure 2.

Figure 3. Word Count and Restrictive Clauses in Australian Federal Legislation, 1975–2014



Source: Produced by the authors using the RDAU1.0 dataset.

Observe a similar trend in the growth rate of regulation between the United States and Australia. A positive correlation suggests that similar factors are driving the patterns of growth in regulation. An obvious factor is correlated economic performance, as related to the similar context of open engagement in a global trading economy and associated global exogenous forces. These may include trade treaties, global environmental treaties, and similar events. The forces of globalization, whether the movement of goods and services, capital, or people, all tend to cause correlated legislative responses. The expansion of the EU and the wave of bilateral trade treaties engendered during the growth phases could explain this. A related factor possibly driving regulatory growth across multiple countries is the political effects of backlash against

globalization, which also tend to cause legislative responses in the form of protectionism, which is often more effectively accomplished through regulation (rather than through direct trade barriers, as before the 1990s). A further factor explaining regulatory growth could be the effects of globalization on international specialization and growing comparative advantage in trade, with regulatory response. This form of economic evolution would show up in correlated patterns of regulation, but across different industries.

3.2. Fixed Costs of Regulation

Comparison of the US and Australian versions of RegData can be used to further explore theories about the economics of regulation. Mulligan and Shleifer (2005) proposed and tested a model using file size data of regulations to make comparisons between US states: they tested a prediction of Demsetz (1967) that because regulation has fixed costs, the extent of regulation will be determined by the size of the market. This predicts that regulatory volumes will scale as a function of population. A larger population can afford more regulation, because it can spread the fixed cost over a larger market.

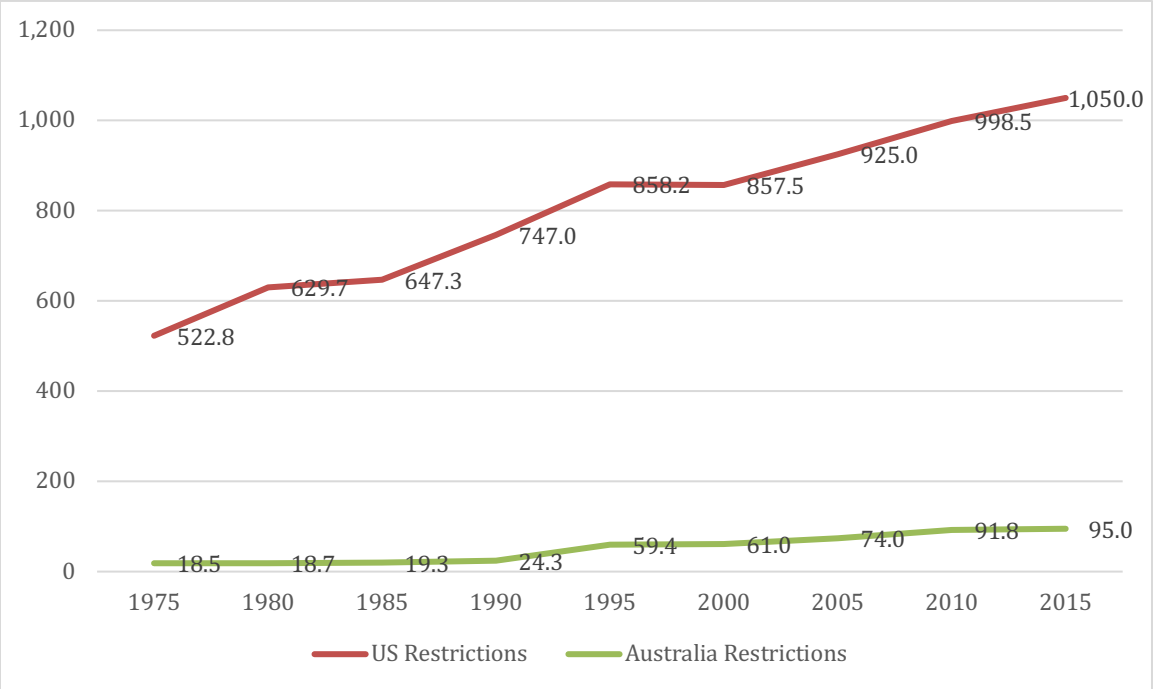
The prediction that Mulligan and Shleifer (2005) tested was whether states with larger populations have more regulations than states with smaller populations.[11] Using file size data on regulations (measured in bytes, and so approximating page numbers and word counts), they found that this prediction was indeed accurate. They then did a cross-country comparison using legal origins as a proxy for fixed costs. Again, they confirmed their hypothesis. So how does their hypothesis hold up under the more rigorous empirical conditions of regulatory constraints?

[11] "The theory predicts that, other things equal, more populous communities should regulate more activities, and do so more intensively. This yields a novel prediction that population is a determinant of the quantity of regulations" (Mulligan and Shleifer 2005: 1447).

Are the US and Australian population sizes consistently correlated with their regulatory volumes? And do the restriction-to-population ratios stay constant over time?

Figures 4, 5, and 6 illustrate what we find when we compare US and Australian regulatory regimes both by page counts and by the more rigorous measures of regulatory restrictions. Figure 4 shows the absolute measures of regulatory restrictions (i.e., the measures uncorrected for population). The US measures are almost an order of magnitude higher than those of Australia, but are similarly upward-trending. This evidence is strongly consistent with Mulligan and Shleifer's (2005) findings, indicating that the United States, which has a much older legislative body (Congress started in the 18th century, whereas the Parliament in Australia dates from 1901) and also a much larger population (318 million versus 23 million in 2014), has a proportionally larger regulatory volume. So this is unsurprising.

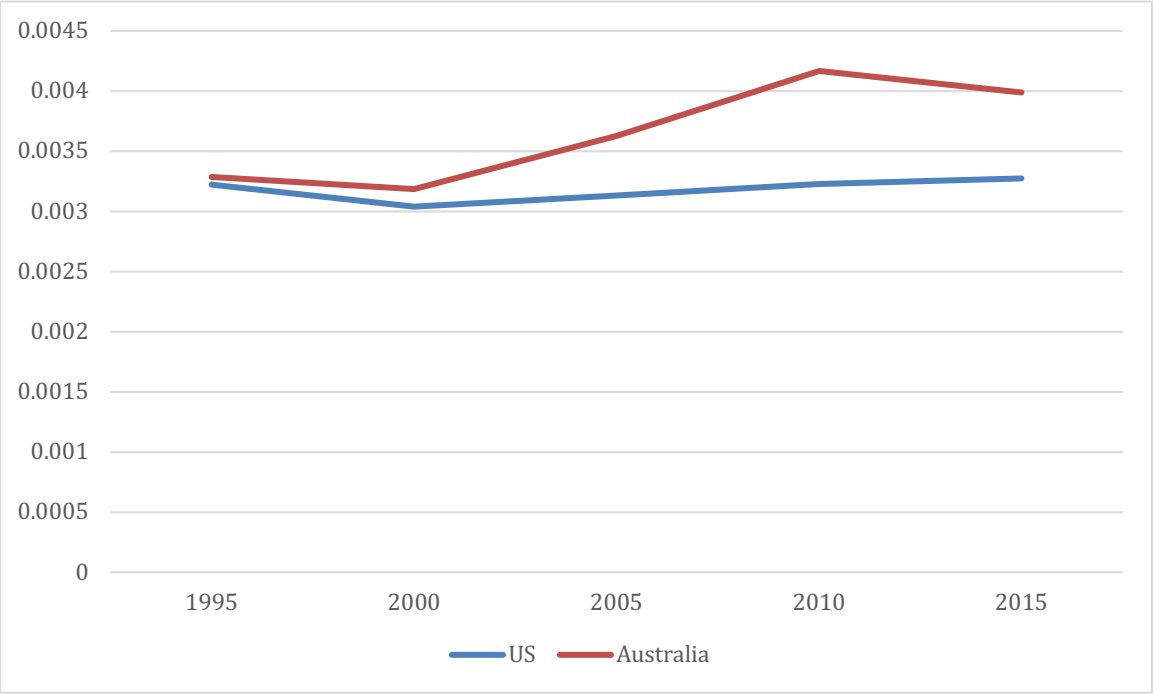
Figure 4. Total US and Australian Federal Regulatory Restrictions (Absolute)



Source: Produced by the authors using the RDAU1.0 dataset and the RegData 2.2 dataset, available at QuantGov, https://quantgov.org/regdata-us/.

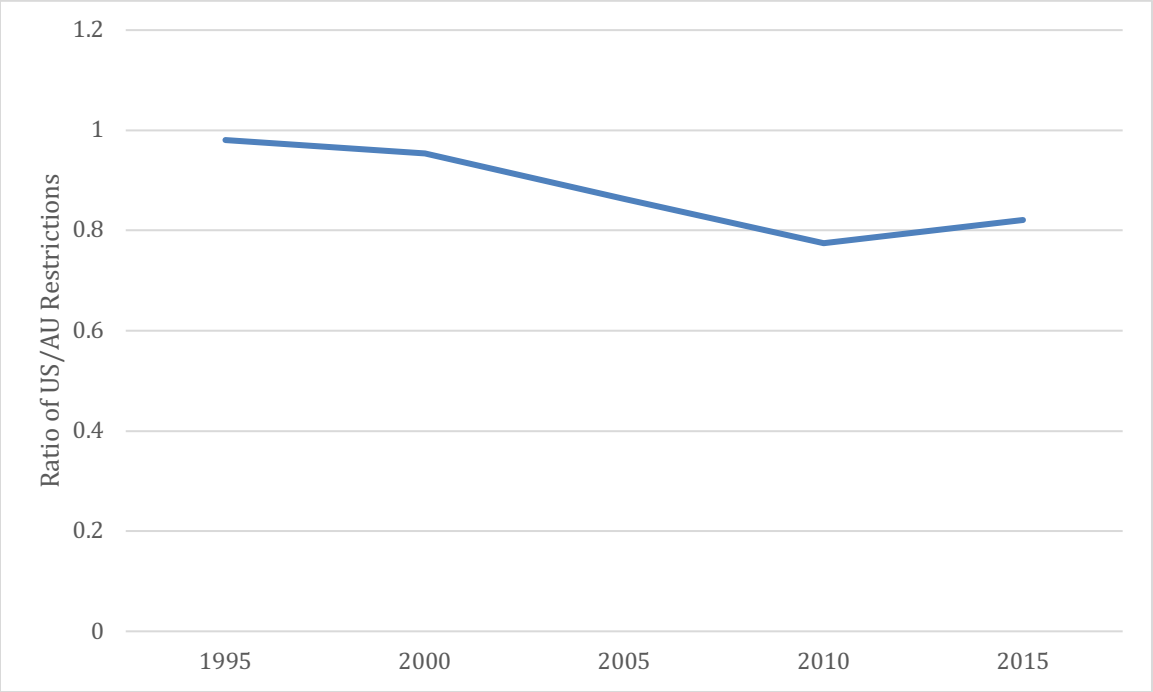
The per capita numbers in figures 5 and 6, in which we focus only on those years for which we are confident about the Australian data, tell a more interesting story. Figure 6 shows that the amount of legislation provided per person is about the same in both countries (i.e., the ratio approximates unity). We leave for further research the question of why it differs from one, and why it appears to vary over time.

Figure 5. US and Australian Federal Regulatory Restrictions per Capita



Source: Produced by the authors using the RDAU1.0 dataset and the RegData 2.2 dataset, available at QuantGov, https://quantgov.org/regdata-us/.

Figure 6. Ratio of US to Australian Per Capita Restrictions

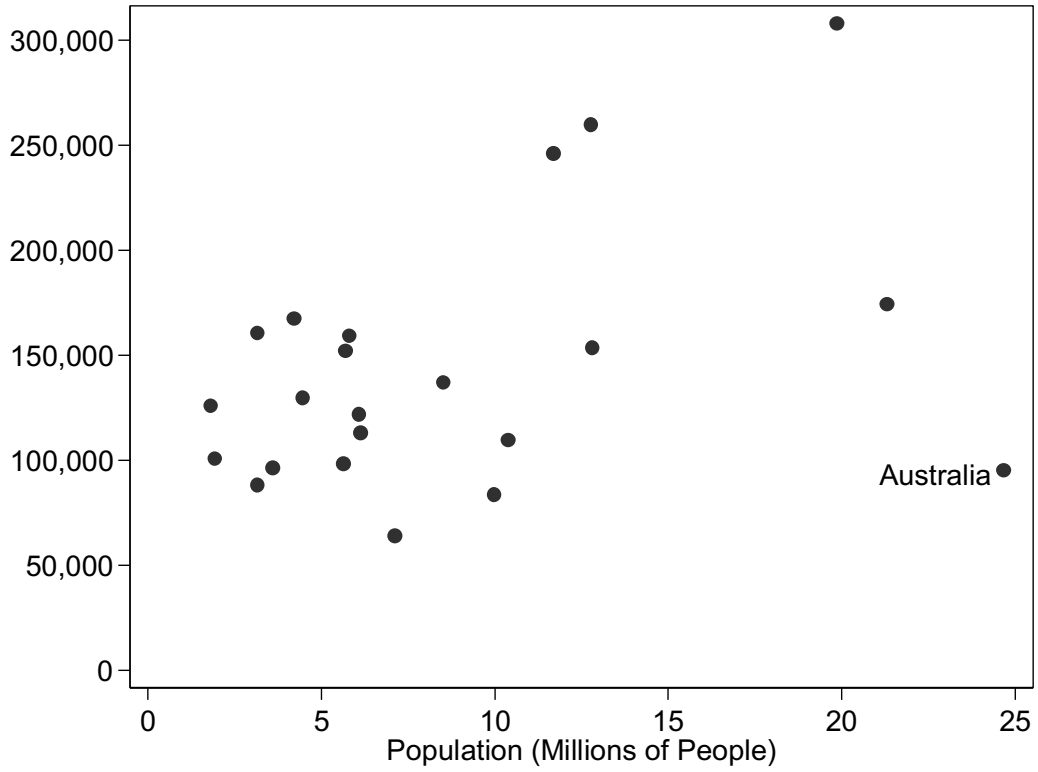


Source: Produced by the authors using the RDAU1.0 dataset and the RegData 2.2 dataset, available at QuantGov, https://quantgov.org/regdata-us/.

These preliminary results, both comparatively and through time, do support Mulligan and Shleifer's (2005) hypothesis about the role of fixed costs in the economics of the supply of regulation.

This hypothesis can be explored further by including regulatory data about various US states. The State RegData project attempts to quantify the amount of regulation in the administrative codes of the US states in a similar way to how the original RegData project quantifies the amount of federal regulation (McLaughlin et al. 2018). Here we compare the 2018 population and regulatory restriction counts of 21 US states and of Australia (we leave US federal regulation out of the comparison because it is such an outlier). Figure 7 and table A1 in the appendix show the comparison and the simple correlation.

Figure 7. Population and Regulatory Restrictions

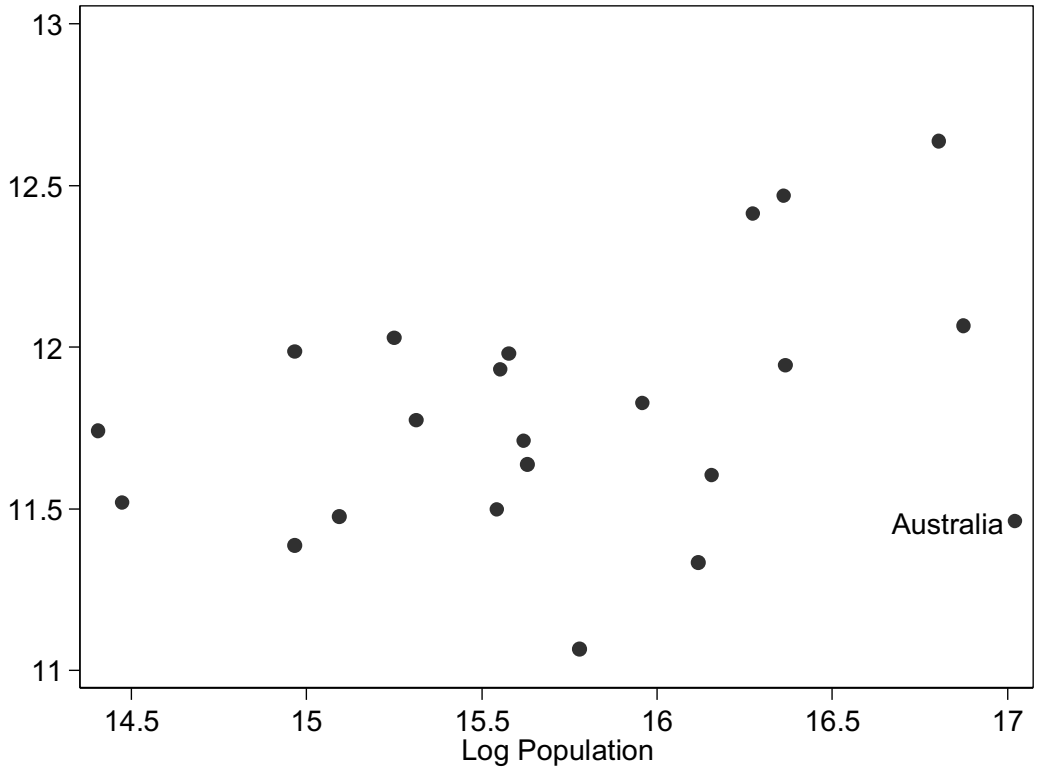


Note: This analysis includes Australia and 21 US states: Arizona, Colorado, Connecticut, Florida, Illinois, Iowa, Kentucky, Maryland, Michigan, Minnesota, Missouri, Nebraska, New York, North Carolina, Ohio, Oregon, Pennsylvania, Utah, Virginia, West Virginia, and Wisconsin.

Sources: McLaughlin, P., Sherouse, O., Francis, D., Nelson, J. (2018), State RegData (dataset), QuantGov (Mercatus Center at George Mason University, Arlington, VA); WorldPopulationReview.com, "US States—Ranked by Population 2018," accessed March 15, 2018, http://worldpopulationreview.com/states/.

There appears to be a positive correlation between population and regulatory restrictions, though a somewhat weak one. Running a basic linear regression with no controls, we find that, on average, an increase in population of 10,000 people is associated with an increase of 39 regulatory restrictions. However, this result is not statistically significant. Since the distribution of both variables is fairly skewed to the right, we next take the log of both population and regulatory restrictions and run a simple linear regression on these variables. Figure 8 and the appendix show the results.

Figure 8. Log Population and Regulatory Restrictions



Note: This analysis includes Australia and 21 US states: Arizona, Colorado, Connecticut, Florida, Illinois, Iowa, Kentucky, Maryland, Michigan, Minnesota, Missouri, Nebraska, New York, North Carolina, Ohio, Oregon, Pennsylvania, Utah, Virginia, West Virginia, and Wisconsin.

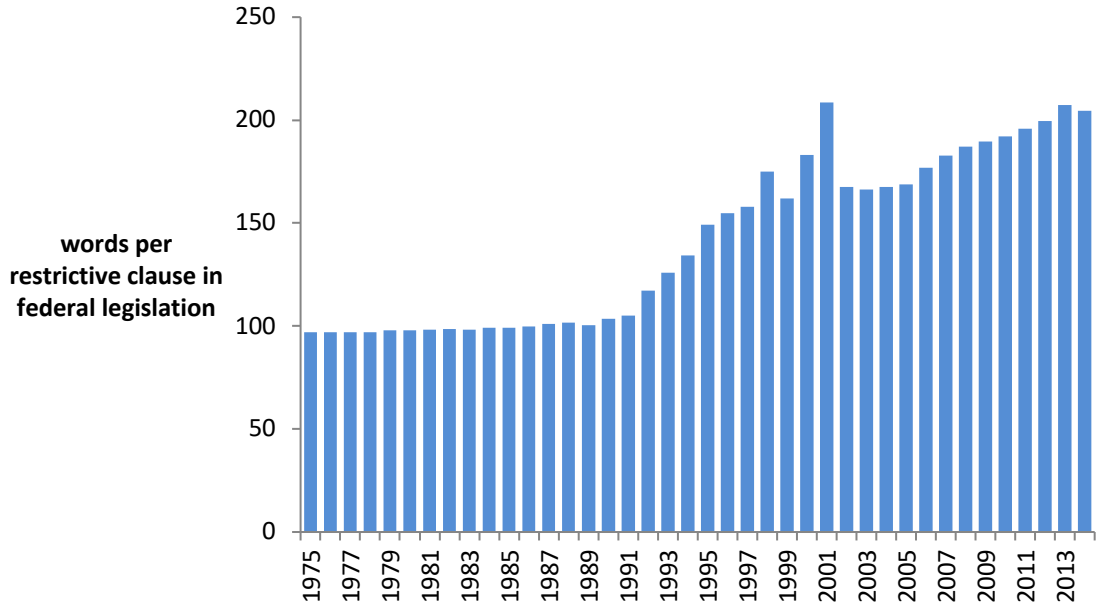
Sources: McLaughlin, P., Sherouse, O., Francis, D., Nelson, J. (2018), State RegData (dataset), QuantGov (Mercatus Center at George Mason University, Arlington, VA); WorldPopulationReview.com, "US States—Ranked by Population 2018," accessed March 15, 2018, http://worldpopulationreview.com/states/.

Logging both variables causes the correlation to weaken slightly, though there is still a seemingly positive relationship. The regression results show that, on average, a 10 percent increase in population is associated with a 1.87 percent increase in regulatory restrictions. Although weak, these correlations are in line with the findings of Mulligan and Schleifer (2005). Future RegData datasets from other countries will help researchers explore this hypothesis further.

3.3. Why Are Regulations Getting More Wordy?

A further empirical observation from the United States and now also the Australian RegData dataset is a curious one that possibly deserves theoretical explanation. It is this: the number of words per restrictive clause is growing. The restrictive clause measure is ideal for capturing this new, possibly stylized, empirical observation. The basic pattern (shown in figure 9) is that words per restrictive clause was more or less constant until 1990 or so. Words per restriction then increased throughout the 1990s, then fell to a new plateau after 2000. But this was not a small shift—rather, the number of words per restrictive clause has doubled since 1975.

Figure 9. Words per Restrictive Clause in Australian Federal Legislation



Source: Produced by the authors using the RDAU1.0 dataset.

The philosopher Harry Frankfurt (2005) has argued that the amount of bullshit in the world is increasing. The data presented in figure 9 provide some legislative-cultural evidence to support this hypothesis. However, a technological or supply-side explanation is that lowered

costs of making and drafting legislation (e.g., cheaper photocopiers, computers, and word processors; availability of internet and digital communication technologies) induce greater supply for equilibrium demand. A demand-side explanation is that voters, especially special-interest voters, can monitor the output of governments and politicians much more intensely now, and that this increases the supply of regulation as volume but not necessarily as intent. A productivity explanation is that regulation entered a period of bureaucratic professionalization as lobbying improved, and the increased wordiness reflects regulators' response. All these explanations posit an efficiency gain, which is manifest in more detailed drafting of legislation and regulation—that is, they suggest that regulation is wordy with a purpose.

Another hypothesis is that increased intensity of the political process, manifesting as increased output of regulation and therefore increased bureaucracy, benefits the agents of that process, namely bureaucrats (Niskanen 1994). The direct incentives to legislators are perhaps small, but the greater quantity and complexity of legislation directly benefits the political class of public servants, whose job it is to oversee, develop, implement, and enforce these regulations. Berg et al. (2005: 15) explain that

> more law means more enforcement, thus more career opportunities. It also creates specialised intellectual capital that legal simplification may reduce or destroy. For example, a complicated tax system tends to increase the future income possibilities of tax officials.

The rise in words per restrictive clause may be prima facie evidence of a systemic rise in the bureaucratic population.

4. Conclusion

RegData is a superior method—compared to page counts, file sizes, and word counts—for studying the incidence of regulation upon an economy and society. For this reason it advances

the scope for analysis of regulation in a particular country. But RegData also creates the potential for a new comparative analysis of regulatory causes and consequences conditioned upon country-level data. This paper has indicated some of the prospects of that line of analysis.

RDAU1.0 provides the first long-run aggregated and disaggregated (by act) time-series measure of the size and strength of regulation in Australia. We hope and expect that this database panel will be of unique value for Australian researchers in a variety of fields. Our ongoing research will next extend the panel to ANZSIC codes, which will benefit industrial economics, including analysis of differential treatment of sectors, and will improve the quality of data before 1997 and also potentially to 1975, which will benefit historical economic analysis. There are many new questions about the Australian economy and its interaction with the regulatory environment that can now be posed along the empirical lines of the RDAU1.0 panel.

But we also expect that the Australian RegData panel will be the second of many subsequent national panels to adopt the RegData method, and will contribute to building a global open-access database of comparable regulatory measures. From this, we hope, will grow a new analytic field of historical and comparative regulatory economics, which will serve as an input into institutional economics, law and economics, public choice economics, growth economics, and development economics.

References

Al-Ubaydli, O., McLaughlin, P. (2014). "RegData: A Numerical Database on Industry-Specific Regulations for All U.S. Industries and Federal Regulations, 1997–2012." Working Paper No. 12-20, Mercatus Center at George Mason University, November. Available at SSRN: http://ssrn.com/abstract=2099814.

Al-Ubaydli, O., McLaughlin, P. (2017). "RegData: A Numerical Database on Industry-Specific Regulations for All U.S. Industries and Federal Regulations, 1997–2012." *Regulation & Governance*, 11(1): 109–23.

Banks, G. (2004). "The Good, the Bad and the Ugly: Economic Perspectives on Regulation in Australia." *Economic Papers*, 23: 22–38.

Banks, G. (2005). "Regulation-Making in Australia: Is It Broke? How Do We Fix It?" Public Lecture Series. The Australian Centre of Regulatory Economics, Australian National University, Canberra.

Berg, C. (2008). *The Growth of Australia's Regulatory State: Ideology, Accountability, and Mega-regulators*. Institute of Public Affairs: Melbourne.

Berg, C., Marohasy, J., Moran, A., Hoggett, J., Warby, M. (2005). "Impact and Outcome of Regulation on the Economy." Occasional Paper, Institute of Public Affairs, Melbourne.

Boettke, P., Coyne, C., Leeson, P. (2013). "Comparative Historical Political Economy." *Journal of Institutional Economics*, 9(3): 285–301.

Boettke, P., Coyne, C., Leeson, P., Sautet, F. (2005). "The New Comparative Political Economy." *Review of Austrian Economics*, 18(3): 281–304.

Butlin, N., Barnard, A., Pincus, J. J. (1982). *Government and Capitalism: Public and Private Choice in Twentieth Century Australia*. Allen & Unwin: Sydney.

Coffey, B., McLaughlin, P., Tollison, R. (2012). "Regulators and Redskins." *Public Choice*, 153(1–2): 191–204.

Coglianese, C. (2002). "Empirical Analysis and Administrative Law." *University of Illinois Law Review*, 2002(4): 1111–1138.

Davidson, A., Elliston, L. (2005). "Regulation and Farm Viability." *Australian Commodities*, 12(1, March): 200–207.

Davis, G., Rahman, J. (2006). "Perspectives on Australia's Productivity Prospects." Working Paper 2006-04, Treasury, Australian Government, Canberra.

Dawson, J., Seater, J. (2013). "Federal Regulation and Aggregate Economic Growth." *Journal of Economic Growth*, 18(2): 137–177.

Demsetz, H. (1967). "Toward a Theory of Property Rights." *American Economic Review*, 57(2): 347–359.

Djankov, S., La Porta, R., Lopez-de-Silanes, F., Shleifer, A. (2002). "The Regulation of Entry." *Quarterly Journal of Economics*, 117: 1–37.

Everett, S. (2006). "Deregulation and Reform of Rail in Australia: Some Emerging Constraints." *Transport Policy*, 13(1): 74–84.

Fels, A. (2004). "Regulating Markets: Marketing Regulation." *Australian Journal of Public Administration*, 63(4): 29–36.

Frankfurt, H. (2005). *On Bullshit*. Princeton University Press: Princeton, NJ.

Glaeser, E., Shleifer, A. (2003). "The Rise of the Regulatory State." *Journal of Economic Literature*, 41: 401–425.

Greenstone, M. (2002). "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures," *Journal of Political Economy*, 110(6): 1175–1219.

Gwartney, James, et al. (2016). *Economic Freedom of the World: 2016 Annual Report*. Fraser Institute: Toronto. September 15.

Hodgson, G. (2015). *Conceptualizing Capitalism: Essence, Institutions, Evolution*. University of Chicago Press: Chicago.

Jaffe, A. B., Peterson, S. R., Portney, P. R., Stavins, R. N. (1995). "Environmental Regulation and the Competitiveness of US Manufacturing: What Does the Evidence Tell Us?" *Journal of Economic literature*, 33(1): 132–163.

Jones, L. (2015). "Cutting Red Tape in Canada: A Regulatory Reform Model for the United States?" Mercatus Research, Mercatus Center at George Mason University.

List, J., Millimet, D., Fredriksson, P., McHone, W. (2003). "Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator." *Review of Economics and Statistics*, 85(4): 944–952.

McLaughlin, P., Greene, R. (2014). "The Unintended Consequences of Federal Regulatory Accumulation." Discussion Paper, Mercatus Center at George Mason University Economic Perspectives series, Arlington, VA.

McLaughlin, P.A. & Sherouse, O. (2019). "RegData 2.2: A Panel Dataset on US Federal Regulations." *Public Choice*, 180: 43–55.

McLaughlin, P., Sherouse, O., Francis, D., Nelson, J. (2018). State RegData (dataset). QuantGov, Mercatus Center at George Mason University, Arlington, VA.

Mulligan, C., Shleifer, A. (2005). "The Extent of the Market and the Supply of Regulation." *Quarterly Journal of Economics*, 120: 1445–1473.

Niskanen, W. (1994). *Bureaucracy and Public Economics*. Edward Elgar: Cheltenham, UK.

Novak, J. (2013). "Australia's Big Government by the Numbers." IPA Report, Institute of Public Affairs.

Pigou, A. (1932). *The Economics of Welfare*. Macmillan: London.

Queensland Competition Authority. (2012). "Measuring and Reducing the Burden of Regulation." Office of Best Practice Regulation: Brisbane, Australia.

Shleifer, A. (2010). "Efficient Regulation." NBER Working Paper No. 15651, National Bureau of Economic Research, Cambridge, MA.

Stigler, G. (1971). "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science*, 2: 3–21.

Thierer, A. (2014). *Permissionless Innovation*. Mercatus Center at George Mason University: Washington, DC.

Appendix: Population and Regulatory Restrictions

Table A1. Population and Regulatory Restrictions in 21 US States and Australia, 2018

	Regulatory Restrictions	
	Level-Level	Log-Log
Population	0.0039	0.1871
	(0.002)	(0.112)
Intercept	109,300	8.8513
	(20,900)	(1.761)

Note: The 21 US states are Arizona, Colorado, Connecticut, Florida, Illinois, Iowa, Kentucky, Maryland, Michigan, Minnesota, Missouri, Nebraska, New York, North Carolina, Ohio, Oregon, Pennsylvania, Utah, Virginia, West Virginia, and Wisconsin.

Sources: McLaughlin, P., Sherouse, O., Francis, D., Nelson, J. (2018), State RegData (dataset), QuantGov (Mercatus Center at George Mason University, Arlington, VA); WorldPopulationReview.com, "US States—Ranked by Population 2018," accessed March 15, 2018, http://worldpopulationreview.com/states/.