5-2017

Pollution and Politicians: The Effect of PM on MPs

Anthony Heyes
University of Ottawa

Nicholas Rivers
University of Ottawa

Brandon Schaufele
Ivey Business School

Pollution and Politicians: The Effect of PM on MPs*

Anthony Heyes† Nicholas Rivers‡ Brandon Schaufele§

May 5, 2017

Abstract

Applying methods of textual and stylometric analysis to all 119,225 speeches made in the Canadian House of Commons between 2006 and 2011, we establish that air pollution reduces the speech quality of Canadian Members of Parliament (MPs). Exposure to fine particulate matter concentrations exceeding 15 $\mu g/m^3$ causes a 3.1 percent reduction in the quality of MPs speech (equivalent to a 3.6 months of education). For more difficult communication tasks the decrement in quality is equivalent to the loss of 6.5 months of schooling. Our design accounts for the potential endogeneity of exposure and controls for many potential confounders including individual fixed effects. Politicians are professional communicators and as such the analysis contributes to our evolving understanding of how pollution exposure impacts the execution of work-relevant skills. Though we are cautious in interpreting the effect as a clean metric for performance, the effect size is around half that established in recent research for workers engaged in physical work tasks. Insofar as the changed speech patterns reflect diminished mental acuity the results make plausible detrimental effects of air pollution on productivity in a wider set of communication-intensive work settings.

Keywords: Air pollution; analysis of speech; non-health impacts; workplace performance

JEL codes: Q52, Q53.

*Funding from the Social Science and Humanities Research Council is gratefully acknowledged. Steve Martin provided excellent research assistance. Special thanks to Bruce C. Young, Head of Parliamentary Publications for Canada's House of Commons. Helpful comments and suggestions were provided by Adam Fremeth, Isaac Holloway, Margaret Loudermilk, Davin Raiha and Jay Shimshack. Errors are ours.

†Department of Economics, University of Ottawa, Faculty of Social Sciences, 120 University, Social Sciences Building, Ottawa, Ontario, Canada, K1N 6N5, 1-613-562-5800 ext. 1135, aheyes@uottawa.ca. Heyes is also part-time Professor of Economics at the University of Sussex.

‡Graduate School of Public and International Affairs and Institute of the Environment, University of Ottawa, Faculty of Social Sciences, 120 University, Social Sciences Building, Ottawa, Ontario, Canada, K1N 6N5, 1-613-562-5800 ext. 4676, nrivers@uottawa.ca.

§Business, Economics and Public Policy, Ivey Business School, University of Western Ontario, 1255 Western Road, London, Ontario, Canada, N6G 0N1, 1-519-200-6671, bschaufele@ivey.uwo.ca.

1 Introduction

Air pollution has been shown to have negative health consequences[1] and decrease cognitive performance.[2] An important recent strand of work has also established a link from short-term pollution exposure to reduced worker productivity. The seminal research relates to manual, outdoor work tasks such that an important challenge for future research is to explore the extent to which these results extend to more highly-skilled and indoor settings.

We make a first foray into linking air quality to how a set of highly-skilled workers perform a *creative* task (public speaking). Workers engaged in creative and cognitively-demanding tasks are often portrayed as the primary drivers of modern economies, so it is important to understand how environmental quality affects, if at all, the work that they do. Measuring the performance of a creative worker is inherently more challenging. Not only is performance normally evaluated along qualitative dimensions, but creative workers often have greater flexibility to reallocate tasks across both time and space and there can be a long and unobserved delays between the period when work is done and when output is observed (the writing of an academic paper provides a perfect example).

Our focus is on a group of professional communicators, namely politicians. Com-

[1]Dockery et al. (1993) demonstrate that American adults in more polluted cities have mortality rates which are 1.26 times greater than in less polluted cities. Pope III et al. (2002) find that moderate increases in fine particulate matter and sulfur dioxide are associated with increased lung cancer, cardiopulmonary disease, and overall mortality in the US, while Xu et al. (1994) and Chen et al. (2013) report similar findings for China. Chay and Greenstone (2003) and Neidell (2004) document a strong causal link between infant mortality and morbidity and pollution, even at low atmospheric concentrations. Internationally, the United Nations (2014) estimates that air pollution is linked to one million premature deaths and one million pre-natal deaths each year, and imposes health costs estimated at approximately 2 percent of GDP in developed economies and 5 percent in developing countries. Beyond the well-documented cardiopulmonary and cancer risks, recent studies have connected ultrafine particulate matter, $PM_{2.5}$, to central nervous system (CNS) function and cognition. Elder et al. (2006) and Oberdörster et al. (2004) show that $PM_{2.5}$ can lead to CNS dysfunction through the circulatory system or even by direct transmission to the brain via breathing.

[2]Lavy et al. (2014a,b) study the implications of air pollution on childrens' test scores in Israel (Bagrut tests). They demonstrate that exposure to $PM_{2.5}$ and carbon monoxide on the day of the test reduces performance.

bining textual and stylometric analyses, we demonstrate that air borne particulate matter causes a statistically significant and substantial reduction in the quality (complexity) of speeches made by Canadian Members of Parliament (MPs).[3]

It is worth contemplating upfront what the patterns that we observe in the data allow us to claim, and what they do not. While speech complexity is not a conventional economic metric, speaking is an important part of the work of a politician (and many other professions) – a central component of his or her day to day job. The mapping from complexity to 'quality' of speech is clearly not a straight-forward one. An MP is a communicator, and as such it would be foolhardy to think of him trying to maximize complexity. However, we might think of him as a rational actor who seeks the 'optimal' level of complexity with which to speak. Then treating the pattern of speech that he delivers on an *un*polluted day as a comparator, systematic deviations from that (in either direction) could sensibly be regarded as involuntary decrements in performance.[4] Since communication is inherently something jointly produced between transmitter and receiver, one potential challenge to our inference is that, if the audience (other MPs) are having their interpretive acuity compromised by pollution, then the speaker may be reoptimising his speech pattern to reflect that. This seems far-fetched, but we have no way to exclude that part of the primal impact falls on the receiver. In that case the decrement to human facility would need to be interpreted as shared, with the polluted air also reducing the capacity of MPs to comprehend complex messages.

Given these and other caveats our findings are suggestive rather than definitive, and we are cautious not to over-interpret the results. However, insofar as the effects extend – in whole or part – to a broader set of communication-intensive (sales, teaching, *etc.*) and creative (writing, design) lines of work, the drag of polluted air on the economy could be substantial.

It is useful to sketch how our results complement recent and emerging evidence

[3]Stylometrics is the statistical analysis of variations in literary style between writers or genres.

[4]Just as we might take the way in which an unintoxicated individual controls a car (as measured by objective metrics such as jerkiness of steering movements, driving distance from vehicle in front, lane positioning) as a benchmark against which to hold the 'performance' of the same individual under the influence of alcohol.

on air pollution and work performance as economists are only starting to understand the rich ways in which economic behavior is influence by short term variations in environmental conditions (e.g., De Silva et al., 2012). Graff Zivin and Neidell (2012) and Chang et al. (2014) provide persuasive evidence that short-term exposure to ozone (O_3) and fine particulate matter ($PM_{2.5}$) reduces the productivity of agricultural laborers engaged in unskilled physical work (fruit picking and handling). Since air pollution inhibits breathing – which is why it is associated with things such as asthma episodes and reduced athletic performance – this is an intuitive result. Li et al. (2015) study textile workers engaged in a repetitive manufacturing task in the severely-polluted city in Hebei province, finding a similar effect for $PM_{2.5}$. While interesting, the direct implications of these studies for understanding the burden of pollution on labor productivity in a developed economy are quite limited. Most work in a modern economy occurs indoors, is not physically-demanding, and is performed in cities with good to very good air quality. Furthermore most *high-value* work in such economies is highly-skilled, cognitively-intensive and often creative.

Two recent studies have made some progress in exploring how far the link from pollution to productivity extends to non-manual work. Chang et al. (2016) find that the number of routine calls processed by a sample of call center employees in China is lower on more polluted days. However, while the call center work takes place indoors it remains low- to semi-skilled (indicative of this is that annual average pay of a call center worker in China is around 2,000 USD, less than half average pay in that country). Interestingly, the reductions in call processing per day uncovered in that study are driven by workers spending more time logged-off on more polluted days, rather than their handling calls more quickly. As such the result is more akin to an intra-day labor supply effect than a 'pure' effect on performance of the sort that we will uncover.[5] Archsmith et al. (2016) find that a panel of Major League

[5]Separately it is worth noting that extrapolation of results derived from studies in China to North American or European settings is hampered by the great differences in prevailing air quality conditions. The setting for our study – Ottawa, Canada – has some of the cleanest air among major cities, as shown in Figure 3 in Appendix C. Days that we will define as "high pollution" would be considered clear in many places (including all Chinese cities). The $PM_{2.5}$ concentration of the *most polluted day* in our study has 70 $\mu g/m^3$ *fewer* of $PM_{2.5}$ than the *average day* in Beijing as reported in Li et al. (2015). The export of results from one place to another would be further exacerbated

Baseball (MLB) umpires make more mistakes in the evaluation of balls and strikes on more polluted days. While the work of an umpire is undoubtedly highly-skilled (reflected in the salaries paid to MLB umpires ranging up to 350,000 USD) it remains a job that is done predominantly outdoors, and while quality-focussed the work task is responsive rather than creative in character. Our objective in this study is to extend this research to a set of professionals in a communication-intensive, creative work setting – namely politicians.

Evaluating the prospective harm of pollution on professionals poses two particular challenges. First, most professional-type workers typically concentrate on quality rather quantity. This makes measuring performance tricky. Second, professionals often have substantial flexibility in how they schedule their work. Someone who feels ill on a given day (perhaps due to high levels of air pollution) may defer work to subsequent days. This makes it difficult to know when work got done. Our setting provides an ideal 'laboratory' within which we can avoid or address these and other challenges.

This study focuses on fine particulate matter ($PM_{2.5}$) in Ottawa, the capital city of Canada. $PM_{2.5}$ is a subset of PM_{10} but displays distinct properties and seasonal variation. Unlike PM_{10} particles, $PM_{2.5}$ particulates are usually too small for visual detection. But $PM_{2.5}$ is longer lived and believed to have larger health implications (cognitive, pulmonary and respiratory effects). Importantly $PM_{2.5}$ can permeate most commercial air filters (Cyrys et al., 2004; Morawska et al., 2001). This makes the effects of fine particulates especially pernicious – employees who are indoors remain exposed to $PM_{2.5}$ at levels similar to those immediately outside the building in which they are working.

Avoidance behavior and endogeneity of pollution exposure offer clear challenges to identification in this area. Our research design credibly avoids these problems by exploiting a situation where the location and timing of work is predefined. Specifically, we apply textual analysis to convert over 100,000 verbal statements made by Canadian MPs from 2006 through 2011 into – among other metrics – speech-specific Flesch-Kincaid grade level indices. This index measures the complexity of an MP's

if there was adaptation of those living in one location to typical local conditions.

speech by the number of years of education needed to accurately understand it. Conditioning on individual fixed effects and other controls, we show that elevated levels of airborne fine particulate matter reduces the complexity of MP speeches. A single high pollution day, defined as daily average $PM_{2.5}$ concentrations greater than $15 \mu g/m^3$, causes a 3.1 percent reduction in contemporaneous speech quality. To put this into perspective, this is equivalent to the removal of 3.6 months of education.

Our central result is identified from within-MP variation in speech. However, it is possible that in addition to the 'within' effect that we estimate, there is also a selection effect. In particular, it is possible that (some) MPs speak relatively less on polluted days. Using cross-sectional variation, we examine this potential reallocation of effort across days and find that individuals whose average speech quality is higher do indeed speak less frequently on high pollution days. As such self-selection combined with inter-temporal reallocation of effort is a second channel through which pollution reduces average, contemporaneous workplace performance.

Finally, we explore potential sources of heterogeneity. Exploiting individual fixed effects for identification, we estimate the average treatment effect on the treated of pollution on speech quality for Members of the Government. Opposition parties, whose role is to hold the Government accountable, have a potential advantage in Parliament. They are able to 'script' or prepare interventions to which the Government representative must respond 'off the cuff', giving the latter a more challenging speaking task. Consistent with the hypothesis that more cognitively-intensive tasks are particularly susceptible to pollution we find that the effect more pronounced for Government than Opposition speakers. Exposure to a high pollution day reduces the average quality of oration by a Government speaker by the equivalent of 6.5 months of schooling.[6]

To the best of our knowledge, speech complexity has not previously been used as a measure of workplace performance. Its application has advantages and disadvantages. A central benefit of the measure is the availability of high frequency data and an accepted set of metrics with which to process it. Still, speech complexity is

[6]We also test for the possibility of heterogeneity of response by age of MP (which is an observable that may proxy for other variables such as health status or experience) finding no effect.

at best a proxy for other high level traits such as creativity, attention and precision, factors on which developed economies increasingly depend for economic growth. We are therefore cautious about over-interpreting these results. Yet, while the match from speech complexity to economic activity is imperfect, our results complement the estimates from Graff Zivin and Neidell (2012), Chang et al. (2014), Li et al. (2015), Chang et al (2016), Archsmith et al (2016), that extend the insight that air pollution damages work performance to a quite different sort of work setting – one that is creative, and communication-intensive – with similar sizes of effect. As such we contribute to a body of emerging evidence that, taken as a whole, point to a pervasive negative causal effect of air pollution on human function.

The rest of the paper is laid out as follows. Section 2 walks through our empirical design. This includes discussing how we measure speech complexity, specifying the conditions needed for credible identification, outlining our dataset and discussing potential biases in our econometric approach. Section 3 then presents our econometric models and results. This section is divided into subsections that examine within individual variation, cross MP sorting and heterogeneity. Section 4 concludes.

2 Research Design

2.1 Conceptual Framework for Estimating a Reduced Form

A large share of MPs' time involves making oral statements in the House of Commons or antechambers. As communication is imperative to their output, we assume MPs target a specific level of speech quality in their verbal communication. We are agnostic about the source of the target – it may be individual-specific or based on party norms. All we assume is that politicians select words and form sentences to articulate (or potentially obfuscate) ideas and that achieving this target, (i.e., communicating ideas), requires expending costly effort. Politicians' speech clarity then is determined by both effort and ability.[7] Our preferred empirical specifications

[7]Similar to Graff Zivin and Neidell (2012), Chang et al. (2014) and Li et al. (2015), air pollution is assumed to influence output via an individual's optimal choice of costly effort. Let the target

use individual fixed effects to control for innate ability, so the effect of pollution can be measured via the direct effect on a reduced-form value function. Empirically, we show that the partial derivative from this reduced-form is negative. But there are few theoretical reasons to expect this particular sign. We do not have information on the precise form of politicians' objective functions. MPs may deliberately target high or low levels of complexity of their speech (i.e., they may expend effort to make their speech more simple or more complex). What matters for empirical identification is that these incentives are orthogonal to realized pollution. Consequently, prior to completing the empirical analysis, it is impossible to sign the partial derivative as reasonable explanations can justify a positive or negative responses.

2.2 Empirical Set-up

Our econometric models and results are presented in section 3. Several important elements of the empirical design are discussed beforehand. This includes the measurement of speech complexity, conditions on the data generating process, an overview of the data and potential biases arising from the empirical models' interpretation.

2.2.1 Measuring Speech Complexity

A key methodological contribution of this study involves quantifying the quality of politicians' verbal outputs. Stylometric analysis is applied in the form of readability indices to convert oral statements into speech-specific measures. We are uninterested

level of speech quality, y, be determined by $y = y(e, a)$, which depends on effort, e, and innate ability, a. The cost of effort is $c = c(e, \alpha)$, where α represents exposure to pollution.

MPs choose optimal effort, $e^* = e(a, \alpha)$, by trading-off the costs and benefits of expending effort up to the point where $(y_e - c_e)\,|_{e=e^*} = 0$. Effort, the MPs' choice variable, is unobservable however. So, using this solution, we define an MP's value function as $Y = y^*(e^*(a, \alpha), a, \alpha)$ and apply the envelope theorem to obtain:

$$\frac{\partial Y}{\partial \alpha} = y_\alpha.$$

This partial derivative gives the direct effect of pollution on speech quality evaluated at a politician's optimal choice of effort. This is the parameter that we actually estimate: the reduced form effect of pollution on speech.

in the subject matter of the speeches, *per se*; rather, we use variables that are derived from these speeches via textual analysis. We start by decomposing each speech into a set of basic constituents such as the number of words, the number of syllables and sentence length. Using methods developed by linguists, these numbers are then recombined via a weighting procedure yielding a scalar that aims to capture how difficult a given text is to understand. We convert every speech made in the Canadian Parliament into a single number that reflects its complexity.

To the best of our knowledge, this approach has not been previously used in economics, but analyses of text and language are becoming more common (e.g., Chen, 2013; Durnev et al., 2013; Baylis, 2015). Popular media also use similar readability metrics, for example, to illustrate the declining complexity of US Presidential States of the Union speeches (Guardian, 2013). Despite its convenience and uniqueness, speech complexity is an imperfect measure of output quality. Individuals may communicate just as effectively irrespective of whether they speak at, say, a grade 11 or 12 level. Caution is therefore warranted when generalizing from our econometric results to economic outcomes. Notwithstanding these caveats, there are reasons to view linguistic complexity as a reliable measure. First, as described, we are not interested in the level of speech complexity but by how much it is affected by pollution. That our preferred index has a natural "grade level" interpretation is merely a convenience, not a necessity. (In fact, we explore other indices and their basic components in robustness checks.) Second, it is important to re-emphasize that politicians, the class of professionals that we investigate, are professional orators. A large share of their job entails making public comments. If their statements are unclear, their message may be misinterpreted and errors propagated. Misstatements may even put their jobs in jeopardy. Therefore, for this particular sample, we believe that it is a reasonable proxy of output quality and provides insight into productivity in occupations that demand high levels of concentration.

Our preferred readability index is the Flesch-Kincaid grade level index. This widely-used metric decomposes a piece of text into counts of sentences, words and syllables and then recombines these counts calculating a single number that reflects the grade level of the text. Specifically, the Flesch-Kincaid grade level index is calculated as:

$$y_{ijt} = 0.39 \left(\frac{\text{total words}_{ijt}}{\text{total sentences}_{ijt}} \right) + 11.8 \left(\frac{\text{total syllables}_{ijt}}{\text{total words}_{ijt}} \right) - 15.59 \tag{1}$$

where y_{ijt} is the Flesch-Kincaid grade level index, which is calculated for speech, j, by MP, i, on a specific date, t. The underlying idea of this particular metric is that a selected section of text should be comprehensible by an individual with an education equivalent to the calculated grade level. Between April 2006 through December 2011 for example, Stephen Harper made 1262 speeches in the House of Commons with a mean Flesch-Kincaid index of 12.2 and a standard deviation of 5.8. This implies that Canada's former Prime Minister's average speech is at roughly a grade 12 level.

Our primary specifications focus on the Flesch-Kincaid grade level index as it is the best known and convenient to interpret. We also use a series of alternatives including: the Coleman-Liau index, the automated readability index, the Flesch reading ease index, and the SMOG index, as well as raw counts of syllables per word and words per sentence. These supplementary measures ensure that any results are not endemic to a specific index. Appendix B presents the formulas for calculating these alternatives.

2.2.2 Conditions on Data Generation Process Needed for Identification

Matching variation in productivity data to variation in air pollution level remains thorny for environmental and labour economics. Graff Zivin and Neidell (2013) emphasize that analysis of pollution, health and productivity must recognize that individual behaviour leads to non-random assignment of exposure. This means that estimating the causal relationship between pollution levels and important economic variables such as productivity is not straightforward. Consequently, we summarize two criteria that must be satisfied in order to justify plausibly exogenous contemporaneous pollution exposure and to estimate the causal effect of pollution on speech complexity.

The first criterion is temporal regularity. Output (speeches in our context) must be generated at regular, *pre-scheduled* intervals, must involve cognitively challenging tasks and be subject to contemporaneous local pollution variation. More importantly though, workers (politicians) must have limited capacity to reallocate tasks across days in response to observed pollution. Both committee meetings and Question Period are scheduled well in advance of realized pollution concentrations. This means that MPs are less able to engage in avoidance behaviour and transfer their work to lower pollution days. MPs are expected to attend and participate in these meetings. We exploit within individual variation, but, more generally, believe that daily air pollution exposure is plausibly exogenous to MPs' expected daily verbal output. Yet, while individual fixed effects enable us to mitigate much of the concern with respect to endogeneity of pollution exposure, we do still observe some avoidance behaviour as MPs with higher average speech complexity speak relatively less on more polluted days.

The second condition is a uniformity criterion. Output must have a relatively stable average level of quality, conditional on individual fixed effects, which is independent of pollution levels and for which there are accepted standards of measurement. In our context, we must consider the audience for politician remarks. The audience of MPs' comments for both House committees and Question Period is essentially constant – and it is largely non-local. MPs in Question Period, for instance, speak to opposing MPs, media observers and to the official record. Speaking to the media and the official record means that politicians are speaking to the public, which is dispersed across the country and not exposed to the same local air quality. Unofficial speeches (those not occurring in Parliament) are more likely to be tailored to time-varying audiences. For example, remarks by Stephen Harper to a kindergarten class will be different from those to the Economic Club. Yet, his audience during Question Period remains virtually unchanged. This is the advantage of focusing on official orations within the Houses of Commons: there are few changes in the audience and we do not expect, to a first approximation, strategic, systematic variation in speech complexity.[8] This average uniformity enables us to

[8]An alternative representation of MPs' formal communication process exists. Rather than orating to people across the country, they may be communicating exclusively with other MPs in the same room – i.e., to listeners who have the same pollution exposure. We believe that the national audience is a better description, yet acknowledge that MPs may instead engage in two-sided local

exploit upward or downward deviations from expected quality.

2.2.3 Data

The primary data required are politician speeches and air pollution concentrations. Transcripts for every oral statement made in the Canadian Parliament are available through a service called Hansard.[9] Hansard, among other things, converts recorded orations into digital text files. Transcription is not verbatim however; texts are altered for clarity. Editors remove familiar verbal ticks such as "um" and "ah" and correct overtly misspoken words. Further, in Canada, Parliamentary business is conducted in two official languages: English and French. MPs speaking in their non-native language are more prone to errors, so Canada's Hansard service applies a more active editorial standard to these cases. As our linguistic indices were calibrated for the English language, all speeches made in French were dropped from our analysis. Finally in 2006, Canada implemented a new recording, indexing and transcription program known as Prism. Prism digitally captures audio. According to Hansard Canada, the move to digital records yields cleaner and clearer recording. It improves indexing (i.e., ensures an accurate match between speech and speaker) and dramatically reduces transcription errors and inconsistencies.[10] Individual transcripts were downloaded and processed, yielding counts for the number of sentences, words, characters and syllables in each speech. These counts are used to construct the Flesch-Kincaid grade level and other indices. Appendix A presents additional information on the dataset construction. The data spans 2006 through 2011 and

communication. Our discussion focuses on speakers, but the alternative interpretation shifts the focus from one where pollution affects productivity through speech to one where it influences others' ability to use inputs (i.e., understand speech). Both interpretations imply that our primary results are meaningful; pollution affects speech complexity, even if MPs, say, use simpler sentences or are more emphatic in response to the listeners' capacity to understand. As we are unable to empirically disentangle listener and speaker effects, we continue to focus on the speaker. Regardless, our main conclusion – that pollution effects mental acuity – holds.

[9] Much of the information on the Hansard services was provided via email and phone between the authors and Bruce Young, the Head of the Parliament of Canada's Hansard service.

[10] Hansard UK, in contrast, uses analogue tape, a system prone to transcription mistakes. The UK attempted to implement a Prism-like recording system, but abandoned the project due to its difficulty and budget overruns.

contain 119,225 speeches made by 488 MPs.[11] Table 1 illustrates the mean and standard deviation of the Flesch-Kincaid grade level corresponds to 11.05 and 7.60.

Canadian daily average $PM_{2.5}$ data are from Canada's National Air Pollution Surveillance Program (Environment and Climate Change Canada, 2016). The monitor we use is approximately 2 kilometers from Parliament Hill (this is the closest monitor to Parliament Hill).[12] $PM_{2.5}$ is small enough to bypass most residential and commercial air filters, implying that individuals are exposed even while remaining indoors.

Figure 3 in Appendix C shows that Ottawa, Canada has some of the best air quality among major international capitals. The mean $PM_{2.5}$ in Ottawa equals 4.86 micro-grams per meter cubed ($\mu g/m^3$). The standard deviation equals 3.91 and the maximum value in the data is $22\mu g/m^3$. This maximum value is $8\mu g/m^3$ less than the Canada-wide air quality standard of $30\mu g/m^3$ (Ontario, 1999) and significantly below the World Health Organization's 24 hour daily mean guideline of $50\mu g/m^3$ (WHO, 2014).

Our preferred models also flexibly control for mean daily temperature and daily precipitation. Weather data from the Ottawa Airport station was retrieved from Environment Canada.[13]

2.2.4 Potential Bias in Model Interpretation

Our data enable us to estimate the relationship between *recorded* speech complexity and *atmospheric* pollution concentrations. This relationship is not exactly what we are interested in. Indeed, we want to estimate the causal effect of individual pollution exposure – not atmospheric concentration – on politicians' actual – not

[11] Our sample also includes Members of the Senate, an appointed body in Canada.

[12] There is only one monitoring station in Ottawa that measures $PM_{2.5}$ concentrations, so we cannot test our results with alternative monitoring stations. However, $PM_{2.5}$ concentrations do not vary significantly across a city because $PM_{2.5}$ remains airborne for an extended period, allowing for efficient mixing. For example, Toronto contains 9 $PM_{2.5}$ monitoring stations, and the pairwise correlation of ambient $PM_{2.5}$ concentrations from these monitoring stations is greater than 0.9 for all monitor pairs and typically greater than 0.95 for monitor pairs.

[13] Coefficients for the weather variables from our preferred specification are presented in Appendix C, Table 7. These show that weather has essentially no effect on measured speech complexity.

recorded – verbal output. The data and interpretation are an imprecise match. This mismatch introduces two sources of "measurement error". First, due to editing and recording errors, our variables measuring recorded speech complexity likely overstate the complexity of the actual oral statements. Likewise, while pollution concentrations are accurately measured, the assignment of pollution exposure to an MP is not. Individual Members may be exposed to differing levels of pollution that depend on residential location or commuting method. Appendix D formalizes how these two sources of measurement error may bias our estimands and how the sources of bias work in same direction to attenuate our estimates. Based on what we know about the mismatch between data and interpretation therefore, our estimated coefficients are biased toward zero and our coefficients should be interpreted as conservative estimates of the true effect. In general however, we expect the total bias to be small, because the causal effect reflected in our primary coefficient is likely small.

3 Results

3.1 Graphical Evidence

Figure 1 fits a kernel-weighted local regression through the data on (residualized) MP speech complexity and ambient pollution concentrations. An Epanechnikov kernel is used. The graph illustrates that there is that there is a nonlinear relationship between pollution and speech complexity. Pollution decreases MP speech quality after a threshold is crossed. This nonlinearity is consistent throughout the literature linking air pollution to productivity. Graff Zivin and Neidell (2012), Li et al. (2015) and Lichter et al. (2015) among others find similar patterns. Importantly however, Figure 1 generalizes a relationship established for manual labour to a broader class of workers, those engaged in mentally, rather than physically, straining activities.

The dependent variable in Figure 1 is the residualized Flesch-Kincaid grade level index for each speech made by Canadian MPs from 2006 to 2011. Residuals are calculated by regressing the Flesch-Kincaid index on parliamentary session, month and day of the week fixed effects and on linear and quadratic daily mean temperature and precipitation variables.

Figure 1 illustrates a nonlinear relationship between speech complexity and pollution. There is a gradual decline in speech complexity until 15 $\mu g/m^3$ of $\mathrm{PM}_{2.5}$, at which point the slope becomes notably steeper. Increasing particulate matter concentrations from 15 to 20 $\mu g/m^3$ yields a reduction in speech complexity of approximately three-quarters of a grade level. The pattern in this figure is nearly identical to the one found for pear packers in Chang et al. (2014).

3.2 Within MP Variation

Econometric Model

Our initial specification identifies a within politician effect of air pollution on speech. We estimate:

$$y_{ijt} = \gamma_i + \nu_s + \mu_d + \rho_m + \phi \cdot Z_t + \boldsymbol{X}_{it}\boldsymbol{\beta} + \varepsilon_{ijt} \tag{2}$$

where y_{ijt} is our measure of complexity (e.g., the Flesch-Kincaid grade level index) for speech, j, from MP, i, on a specific date, t. ν_s, μ_d and ρ_m are session, day of the week and month fixed effects, respectively. MPs may experience common systematic variation in their mental acuity (e.g., end of the week exhaustion), so these coefficients control for these factors. γ_i is an individual, MP-level fixed effect. Any effect of pollution on speech complexity is therefore identified within individuals, not across MPs. In essence, identification depends on differences from MPs' average personal speech level on high and low pollution days. We assume that speeches are targeted at non-local, nation-wide populations (i.e., the uniformity criterion discussed in section 2.2.2 holds), enabling us to assert that these individual averages are not strategically manipulated according to observed pollution. Daily newscasts show excerpts from Parliament and people living in other regions have different realizations of pollution, so this assumption is viewed as mild. This specification also includes a series of time-varying, weather controls in $\boldsymbol{X}_{it}$. The speech-specific error term is given by ε_{ijt}.

Z_t is dummy variable reflecting a "high pollution" day and is defined as:

$$Z_t = I(PM_{2.5} \geq 15) \tag{3}$$

The pollution threshold of $15\mu g/m^3$ for $PM_{2.5}$ is based on the literature and informed by Figure 1.[14] Chang et al. (2014) use thresholds at 15, 20, and $25\mu g/m^3$ for $PM_{2.5}$ in their study. Because Ottawa is relatively unpolluted, we cannot test these higher thresholds. The parameter of interest in (2) is ϕ. This represents the change in an individual MP's speech complexity when air pollution is elevated compared to that same individual's speech complexity at lower levels of pollution. The definition of Z_t in (3) implies that our regression equation, (2), is linear in treatment (high pollution) but nonlinear in pollution exposure. We also provide supplementary results from a linear and log-linear model. Table 6 in Appendix C presents results where $PM_{2.5}$ enters continuously rather than as a binary variable, as well as results derived with a greater number of bins. The results are qualitatively similar to those presented in the main text.

Results

Table 2 presents our main results. This table has six columns. Overall, we observe a robust, unambiguously negative effect of $PM_{2.5}$ on politician speech complexity. We successively add parameters to ensure the stability of estimates. Even though we do not think it is an issue in this context, fixed effects can exacerbate bias in regressions that have time-varying omitted variables. So, columns (1) and (4) exclude all controls and fixed effects. Column (1) shows that being exposed to an average daily $PM_{2.5}$ level greater than 15 $\mu g/m^3$ reduces MPs' average level of speech complexity by 0.40 grade levels (4.8 months of education). Column (4), using a logged Flesch-Kincaid index, corroborates this estimate. It shows that exposure to high daily $PM_{2.5}$ levels yields a 3.5 percent reduction in speech complexity. Columns (2) and (5) include weather controls and time fixed effects. They display estimates of -0.31 and -3.2 percent, respectively. Our preferred models are columns (3) and (6).

[14] There are 5,288 observations that are greater than or equal to $15\mu g/m^3$ of $PM_{2.5}$.

These regressions contain the full set of MP fixed effects in additional to time and weather controls. Within individual variation in pollution exposure is exploited for identification. Column (3) shows that exposure to daily average $PM_{2.5}$ pollution greater than 15 $\mu g/m^3$ reduces the average MP's speech grade level from 11.0 to 10.7. Column (6) illustrates that this is a 3.1 percent decrease. These models match the pattern illustrated in Figure 1.

Several comments are warranted on these results. First, the magnitude of this effect is slightly smaller than the 5.5 percent found by Graff Zivin and Neidell (2012) and the 6.6 percent estimated by Chang et al. (2014). Still, the prospective implications are large. Unlike the prior research, this study focuses on cognitively challenging jobs. This adds meaningful credibility to these prior estimates and supports their generalizability to the wider economy. Even the relatively small effects such as those we report imply huge economic consequences once aggregated over economic activities. Second, even though our point estimates for Ottawa are small, they are statistically significantly different from zero at conventional levels and are credibly identified based on time series variation within individual speakers. Finally, it is worth re-emphasizing that the air quality in Ottawa, Canada is among the best in the world. Days that we define as "high pollution" would be considered clear in many cities. The $PM_{2.5}$ concentration of the *most polluted day* in our study has 70 $\mu g/m^3$ fewer of $PM_{2.5}$ than the *average day* in Beijing as reported in Li et al. (2015).

Results for Alternative Speech Complexity Metrics

The Flesch-Kincaid grade level index has intuitive appeal. Evaluating the comprehensibility of text by years of education is convenient. People understand what ten or twelve years of schooling implies. To ensure these results are not driven by unique features of this particular index however, we re-estimate (2) using six alternative dependent variables. Each of these dependent variables is either another commonly used readability index or one of its components. Formulas for each index are contained in Appendix B.

Table 3 presents six models that mimic the results from Table 2. All models contain the full suite of fixed effects and weather controls. Column (1) shows results for the Coleman-Liau index. With the Coleman-Liau, elevated pollution leads to a 0.15 point reduction in speech quality in Ottawa. Column (2), using the automated readability index, finds similar results – MPs' speech complexity declines by 0.36 points only highly polluted days. The Flesch reading ease score, as presented in (3), has the opposite interpretation from the other indices – higher values indicate lower complexity. Again we observe that $PM_{2.5}$ reduces the within individual complexity of MP speech. Next is the SMOG index. Column (4) corroborates the other results: $PM_{2.5}$ reduces the speech complexity of Canadian MPs. Note that in strict sense the SMOG index is only valid when applied to longer texts than some of those here, so we interpret this last result cautiously.[15]

Finally, we investigate key components of these indices: syllables per word and words per sentences. Column (5) shows that $PM_{2.5}$ reduces by 0.01 the average syllables per word, while column (6) shows that they speak 0.48 fewer words per sentence. Hence, MPs use shorter words and shorter sentences when exposed to pollution, a result that suggests some cognitive impairment.

3.3 Pollution and the Contemporaneous Selection of Speakers

Within individual variation enables identification of the individual effect of pollution on speech complexity. MPs may also select who speaks on a given day according to observed pollution. This selection effect is independently interesting as it demonstrates an alternative channel through which pollution can reduce workplace performance – i.e., via the reallocation of tasks from more to less productive workers. More senior MPs, say, may both be better average speakers and more susceptible to pollution-related health problems such as respiratory inflammation. If pollution levels are elevated, we may observe that on polluted days average speech complexity declines even though no within individual effect is identified. Cross sectional

[15]Having acknowledged that, the index acronym – SMOG – made it hard to exclude from a paper on this topic!

variation between MPs, in other words, may mask the effect of pollution on speech complexity as some individuals may engage in avoidance behaviour.

We explore age-related heterogeneity in the next subsection. First however, we explore a more general approach to estimating the potential selection of MP speeches across days. We employ a model similar in spirit to the correlated random effects model of Wooldridge (2009) and Mundlak (1978).[16] We simultaneously examine within MP effects as in Table 2 and cross-MP variation. Cross-sectional variation arises because MPs speeches are dispersed across time – i.e., even if they are present in Parliament, not every MP speaks every day. We specify:

$$Y_{ijt} = \nu_s + \mu_d + \rho_m + \psi_1(Z_t - \bar{Z}^i) + \psi_2\bar{Z}^i + \boldsymbol{X}_{it}\boldsymbol{\beta} + \varepsilon_{ijt} \tag{4}$$

The fixed effects, ν_s, μ_d and ρ_m, and control variables, $\boldsymbol{X}_{it}\boldsymbol{\beta}$, are identical to (2). Z_t defined in two ways. In one model it is as in (3). A second model uses continuous daily average pollution concentration. What distinguishes (4) from (2) is $\bar{Z}^i$. $\bar{Z}^i$ is the average pollution exposure of MP, i, over days, t, on which she spoke. This variable captures the mean cross MP variation is speech complexity explained by pollution. The estimands in this specification are ψ_1 and ψ_2. Wooldridge (2009) shows that, even though the model is estimated via random effects, conditional our assumption on MP-level variation, the coefficient on within effect, ψ_1, is equivalent to the standard fixed effects estimate. ψ_2 then provides additional information on the between or cross-sectional effect. That is, it provides information on a psuedo-selection effect: how the composition of MPs – as defined by their average level of

[16] As a supplementary check, to ensure that our within individual results are not driven by selection on unobservables, we also calculate Altonji et al.'s (2005) selection on observables to selection on unobservables ratio. This statistic uses selection on observables to determine the how large the bias created by selection on unobservables needs to be to fully explain our estimated coefficient. We calculate the statistic based on model (3) from Table 2 (i.e., including MP fixed effects). The calculation shows that the ratio of selection on observables to selection on unobservables is 92.9, much larger than rule of thumb of one for a robust estimate. In other words, selection on unobservables would need to be 92.9 times larger than selection on observables to fully explain our effect size. While it is not possible to assign economic meaning to this ratio, it strongly suggests that time-varying, MP-specific variables are not biasing our coefficients in Table 2 and suggests that avoidance behaviour is likely limited.

pollution exposure on days which they actually speak – changes as pollution changes.

Table 4 presents results from two models. The model in (1) contains two coefficients and includes $PM_{2.5}$ as a continuous regressor. The within MP effect shows that an increase of 10 $\mu g/m^3$ of $PM_{2.5}$ reduces speech complexity by a statistically significant -0.21 grade levels. This is equivalent to two and a half months of schooling. The cross MP parameter is not statistically distinguishable from zero at conventional levels, but does suggest the existence of some selection with individuals categorized by lower average speech complexity speaking more on more polluted days. Column (2) is similar to model (3) in Table 2. It includes the nonlinear effect of pollution exposure on the Flesch-Kincaid index. The within individual coefficient, which equals -0.40, is similar in magnitude to the previous results. This states that exposure to a high pollution day reduces speech quality by an amount equal to four months of education when compared to the same MP's speech on a low pollution day. The cross MP effect in (2), like in (1), is not statistically significantly different from zero but does suggest that, on average, individuals with lower mean speech complexity make statements on polluted days.

We caution against over-interpreting the cross-sectional coefficients due to their imprecision. Still, while the the standard errors are large, these estimates qualitatively suggest that oral statements made in the Canadian House of Commons are simpler on days with high pollution. This weak evidence accords with the notion that (some) individuals engage in avoidance behaviour when exposed to pollution.

3.4 Heterogeneity

Next, we investigate three sources of potential heterogeneity, which refine our understanding of the mechanism through which pollution might alter MP performance. The econometric model and results are discussed next. First, we provide a brief overview of the models.

To start, some MP speeches are read verbatim from a prepared text while others are spoken in the moment. Information is unavailable to distinguish between these two sets.[17] Situations exist however, in which we may expect fewer scripted statements. Question Period is one example. Question Period allows Opposition members to grill the Government. On occasion, Question Period devolves into a free-for-all where comments are made off-the-cuff and MPs must respond without notes. Pollution may have a larger effect during this melee. Empirically we do not detect any meaningful difference of pollution on speech complexity during Question Period compared with other Parliamentary sessions.

Second, Opposition Members typically initiate questions in the House of Commons. This means that they are better able to script their remarks. Members of the Government, in contrast, are forced to respond to these comments with less preparation. We do find sizable implications for the Government-pollution interaction, suggesting that Members' impromptu speeches are more influenced by pollution than average.

Finally, individuals with different birth years may react differently to pollution exposure. While we acknowledge that age is an imperfect proxy, it is observable and it may capture experience or health status. No notable results are found for age heterogeneity however.

Econometric Model and Results

We explore heterogeneity by estimating the average treatment effect on the treated of air pollution on speech quality. We specify:

$$y_{ijkt} = \gamma_i + \nu_s + \mu_d + \rho_m + \phi_3 \cdot Z_t + \phi_4 \cdot D_k + \phi_5 \cdot Z_t * D_k + \boldsymbol{X}_{it}\boldsymbol{\beta} + \varepsilon_{ijkt} \qquad (5)$$

where the dependent variable, fixed effects and weather controls are as in (2). Z_t is as in (3). D_k is a dummy variable that equals one if: (i) a speech is made in Question Period, (ii) a speech is made by a member of the governing Conservative Party,[18] or (iii) whether an MP was born after a specific year.

[17] Assuming that only speeches spoken in the moment are affected by pollution, our estimates provide a lower bound estimate for the effect of pollution on air quality.

[18] The Conservative Party was in power throughout our sample; as a result, we are unable to disentangle a party effect from a government effect.

Table 5 presents results for five formulations of (5). Column (1) considers whether speeches are made during Question Period. Column (2) examines heterogeneity in the speaker according to their membership in the governing party. The final three columns then explore heterogeneity via birth year.

Question Period does not seem to meaningfully influence speech quality. The direct effect of Question Period in column (1) is -0.34, which is roughly the same magnitude as the main effect on pollution. This estimate is imprecise however. The interaction between a high pollution day and Question Period has even wider standard errors, a positive sign on the coefficient and a smaller magnitude. All told, Question Period has low explanatory power and does not seem to influence MP speech quality.

We contrast these results with column (2). Column (2) estimates the average treatment effect on the treated for members of the governing Conservative Party. Being both a Member of the Government and exposed to a high pollution day decreases speech complexity of 0.53 grade levels or 6.5 months. This effect is statistically significant. In fact, including the Government-pollution interaction attenuates the main pollution coefficient by more than a factor of five. This suggests that pollution may dull MPs' capacity to think quickly or to devise unrehearsed statements.

Finally, columns (3), (4) and (5) show that age has minimal explanatory value. Column (3) looks at MPs born in 1970 or later. This includes 7.3 percent of the sample. While the point estimate is negative, it is small and imprecisely estimated. Column (4) increases the cutoff to MPs who were born after 1959, so captures 31.9 percent MPs. Again, a small, imprecise point estimate is shown. Finally, column (5) adds another decade, by looking at Members born after 1949 (68.3 percent of the sample). Here the point estimate is positive and larger, but still statistically insignificant. Overall, the interaction between age and exposure to a high pollution day does not appear to be important.

4 Conclusion

An important recent strand of research has identified a link from short term air pollution to the performance of workers engaged in physical (Graff Zivin and Neidell (2012)), non-physical but low-skilled (Chang et al. (2016)) and high-skilled but responsive (Archsmith et al. (2016)) work tasks. In this paper we push the boundary further by providing evidence of the impact of polluted air on the same-day performance of a group of professionals engaged in a creative, cognitively-demanding, communication-intensive task. Taken together this group of studies builds an increasingly persuasive case that polluted air is inhibiting performance across a broad swathe of activities.

It is evident that the work done by an MP in the House of Commons is idiosyncratic in character, leading us to be cautious in extrapolating to possible effects across other work settings. However, while the work tasks are indeed idiosyncratic, the same could be said of almost any other high-skilled occupation (teacher, barrister, air traffic controller). In fact idiosyncratic or 'specialist' work more or less defines such occupations. Just as fruit-picking is an idiosyncratic activity from which we seek to extrapolate to a broader set of physically-oriented tasks, the work of an MP might also be taken as an exemplar of creative and communication-intensive lines of work. Understanding the extent to which the evidence-base has to be built profession-by-profession, or the extent to which we can export results derived in one work setting to other sorts of employment requiring similar skill-sets, should be a central ambition of future research.

The analyses have notable policy implications. Canadian and American environmental policies are screened using cost-benefit tests prior to implementation. The benefits side of existing cost-benefit studies are populated almost exclusively by health outcomes, and none – to the best of our knowledge – have sought to account for the beneficial effect that air quality improvements can be expected to have on labour productivity. The decrements in productivity identified are large enough that, if replicated across broader parts of the economy, could plausibly compete in size with the health effects. This body of research, by evidencing a currently uncounted benefit, implies that regulations as currently promulgated will be insufficiently stringent, perhaps substantially so.

References

Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124(4):1729–1772.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of political economy*, 113(1):151–184.

Baylis, P. (2015). Temperature and temperament: Evidence from a billion tweets. Working paper #265, Energy Institute, Haas School of Business, University of California, Berkeley.

Bharadwaj, P., Gibson, M., Zivin, J. G., and Neilson, C. A. (2014). Gray matters: Fetal pollution exposure and human capital formation. Technical report, National Bureau of Economic Research.

Bound, J., Brown, C., Duncan, G. J., and Rodgers, W. L. (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics*, pages 345–368.

Chang, T., Zivin, J. S. G., Gross, T., and Neidell, M. J. (2014). Particulate pollution and the productivity of pear packers. Working paper number 19944, National Bureau of Economic Research.

Chay, K. Y. and Greenstone, M. (2003). Air quality, infant mortality, and the clean air act of 1970. Working paper number 10053, National Bureau of Economic Research.

Chen, M. K. (2013). The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *The American Economic Review*, 103(2):690–731.

Chen, Y., Ebenstein, A., Greenstone, M., and Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from chinaâ€ï¿$\frac{1}{2}$s huai river policy. *Proceedings of the National Academy of Sciences*, 110(32):12936–12941.

Cyrys, J., Pitz, M., Bischof, W., Wichmann, E., Heinrich, J., et al. (2004). Relationship between indoor and outdoor levels of fine particle mass, particle number concentrations and black smoke under different ventilation conditions. *Journal of Exposure Science and Environmental Epidemiology*, 14(4):275–283.

De Silva, D. G., Pownall, R. A. J. and Wolk, L. (2012). Does the sun 'shine' on art prices? *Journal of Economic Behavior & Orginization*, 82(1):167–178.

Deryugina, T. and Hsiang, S. M. (2014). Does the environment still matter? daily temperature and income in the united states. Technical report, National Bureau of Economic Research.

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris Jr, B. G., and Speizer, F. E. (1993). An association between air pollution and mortality in six us cities. *New England journal of medicine*, 329(24):1753–1759.

Durnev, A., Fauver, L., and Gupta, N. (2013). When talk is not cheap: The corporate value of political rhetoric.

Elder, A., Gelein, R., Silva, V., Feikert, T., Opanashuk, L., Carter, J., Potter, R., Maynard, A., Ito, Y., Finkelstein, J., et al. (2006). Translocation of inhaled ultrafine manganese oxide particles to the central nervous system. *Environmental health perspectives*, pages 1172–1178.

Environment and Climate Change Canada (2016). Fact sheet: Ambient (outdoor) air quality and health. Technical report, "http://www.ec.gc.ca/rnspa-naps/" (accessed April 11, 2016).

Graff Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102:3652–3673.

Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51:689–730.

Guardian, N. (2013). The state of our union is dumber: How the linguistic standard of the presidential address has declined. Technical report, "http://www.theguardian.com/world/interactive/2013/feb/12/state-of-the-union-reading-level" (accessed April 6, 2016).

Hanna, R. and Oliva, P. (2011). The effect of pollution on labor supply: Evidence from a natural experiment in mexico city. Technical report, National Bureau of Economic Research.

Lavy, V., Ebenstein, A., and Roth, S. (2014a). The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation. Technical report, National Bureau of Economic Research.

Lavy, V., Ebenstein, A., and Roth, S. (2014b). The long run human capital and economic consequences of high-stakes examinations. Technical report, National Bureau of Economic Research.

Li, T., Liu, H., and Salvo, A. (2015). Severe air pollution and labor productivity. *Available at SSRN*.

Lichter, A., Pestel, N., and Sommer, E. (2015). Productivity effects of air pollution: Evidence from professional soccer.

Morawska, L., He, C., Hitchins, J., Gilbert, D., and Parappukkaran, S. (2001). The relationship between indoor and outdoor airborne particles in the residential environment. *Atmospheric Environment*, 35(20):3463–3473.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, pages 69–85.

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, 23(6):1209–1236.

Oberdörster, G., Sharp, Z., Atudorei, V., Elder, A., Gelein, R., Kreyling, W., and Cox, C. (2004). Translocation of inhaled ultrafine particles to the brain. *Inhalation Toxicology*, 16(6-7):437–445.

Ontario, G. o. (1999). Ontario and the canada-wide standards for particulate matter and ground-level ozone. Technical report, Ministry of Environment.

Pope III, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., and Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama*, 287(9):1132–1141.

Power, M. C., Weisskopf, M. G., Alexeeff, S. E., Coull, B. A., Spiro III, A., and Schwartz, J. (2011). Traffic-related air pollution and cognitive function in a cohort of older men. *Environmental Health Perspectives*, 119(5):682.

WHO, W. H. O. (2014). Fact sheet: Ambient (outdoor) air quality and health. Technical report, World Health Organization: “http://www.who.int/mediacentre/factsheets/fs313/en/” (accessed March 28, 2016).

Williams, R. C. (2002). Environmental tax interactions when pollution affects health or productivity. *Journal of Environmental Economics and Management*, 44(2):261–270.

Wooldridge, J. M. (2009). Correlated random effects models with unbalanced panels. *Manuscript (version July 2009) Michigan State University.*

Xu, X., Gao, J., Gao, J., and Chen, Y. (1994). Air pollution and daily mortality in residential areas of Beijing, China. *Archives of Environmental Health*, 49(4):216–222.

5 Figures

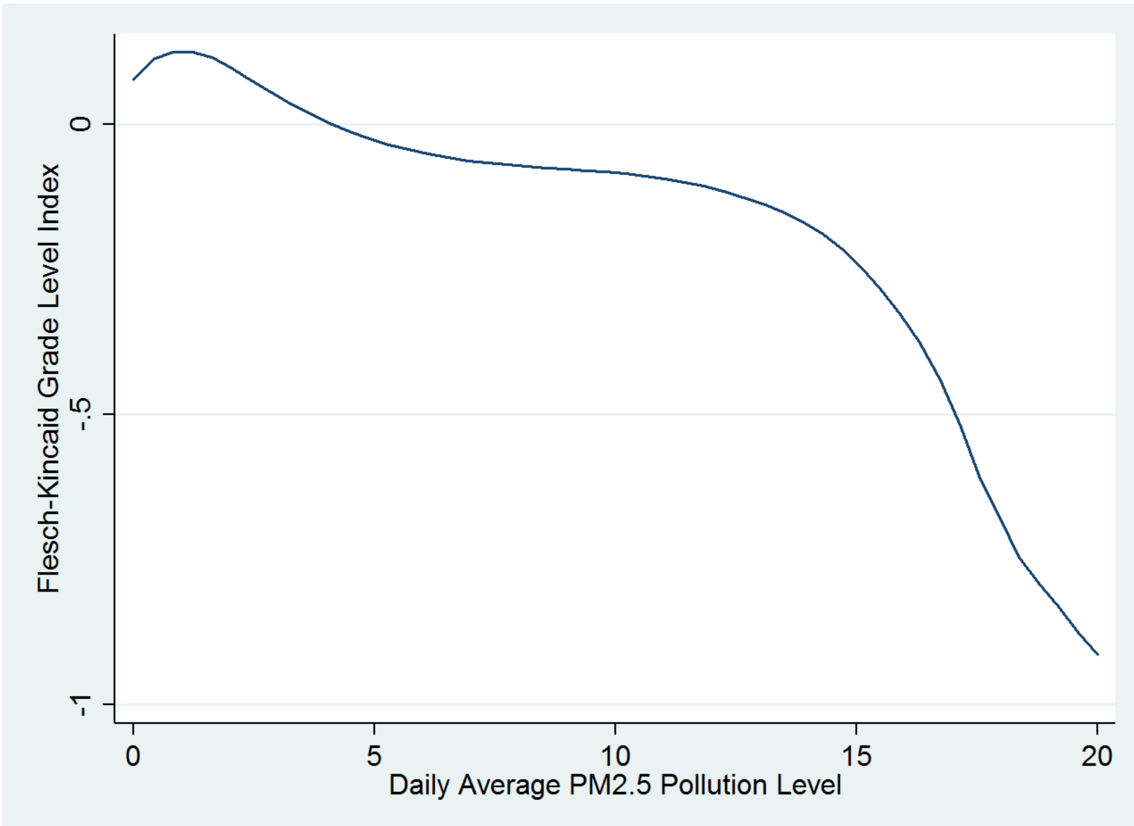

Figure 1: Non-linear effect of daily average $PM_{2.5}$ levels of politician speech complexity. The curve is generated using a kernel-weighted (Epanechnikov) local regression fit through residual Flesch-Kincaid index for all MP speeches in the data set. Residuals are generated from a regression of speech complexity on temperature, temperature squared, precipitation, precipitation squared, and MP, parliamentary session, and day of week fixed effects. Bandwidth = 6 $\mu g / m^3$

6 Tables

Table 1: Summary Statistics

	Mean	Standard Deviation
$PM_{2.5}$ concentration ($\mu g/m^3$)	4.86	3.91
Flesch-Kincaid grade level	11.05	7.60
Coleman-Liau index	10.92	4.51
Automated readability index	10.97	9.61
Flesch reading ease score	55.05	31.66
SMOG index	12.38	4.90
Syllables per word	1.53	0.28
Words per sentence	22.00	15.52
Temperature (°C)	5.85	10.05
Precipitation (mm)	6.21	23.50
Number of MPs	489	
Number of speeches	119,225	

Data includes all speeches made in the Canadian House of Commons by Members of Parliament between 2006 and 2011. Daily mean pollutant concentration is derived from the "Ottawa Downtown" monitoring station at Rideau and Wurtemberg (NAPS id 60104). Weather data is obtained from the Ottawa Airport weather station.

Table 2: Effect of Elevated Particulate Matter Pollution Concentrations on the Speech Complexity of Canadian Members of Parliament

	Flesch-Kincaid Index			log(*Flesch-Kincaid Index*)		
	(1)	(2)	(3)	(4)	(5)	(6)
$I(PM_{2.5} \geq 15\ {}^{\mu g}/_{m^3})$	-0.401	-0.311	-0.303	-0.035	-0.032	-0.031
	(0.145)	(0.150)	(0.151)	(0.017)	(0.018)	(0.018)
Weather controls		✓	✓		✓	✓
Day-of-week fixed effects		✓	✓		✓	✓
Month fixed effects		✓	✓		✓	✓
MP fixed effects			✓			✓
Number of MPs	488	488	488	480	480	480
Observations	119,225	119,225	119,225	110,913	110,913	110,913

Values in parentheses are standard errors clustered on individual MPs.

Weather controls include temperature, precipitation, and their squares.

Table 3: Effect of Elevated Particulate Matter Pollution on MP Speech Complexity, Alternative Indices

	Coleman Liau (1)	Automated Readability (2)	Flesch Readability (3)	SMOG Index (4)	Syllables per Word (5)	Words per Sentence (6)
$I(PM_{2.5} \geq 15\ {}^{\mu}/_{m^3})$	-0.149	-0.357	1.328	-0.140	-0.010	-0.475
	(0.087)	(0.163)	(0.770)	(0.095)	(0.008)	(0.250)
Number of MPs	488	488	488	488	488	488
Observations	119,225	119,225	119,225	119,225	119,225	119,225

Values in parentheses are bootstrapped standard errors clustered on individual MPs.

All models contain fixed effects for day of week, month, and MP, as well as weather controls.

Table 4: Comparison of Time Series and Cross Sectional Effects of Pollution on MP Speech

	$PM_{2.5}$ (1)	$I(PM_{2.5} \geq 15)$ (2)
Within MP effect	-0.021	-0.406
	(0.009)	(0.139)
Cross MP effect	-0.106	-1.903
	(0.080)	(1.153)
Number of MPs	488	488
Observations	119,225	119,225

Values in parentheses are bootstrapped standard errors.

Table 5: Heterogeneous Effects of Particulate Matter Pollution on MP Speech Complexity

	(1)	(2)	(3)	(4)	(5)
$I(PM_{2.5} \geq 15)$	-0.361	-0.062	-0.346	-0.399	-0.584
	(0.197)	(0.180)	(0.189)	(0.267)	(0.364)
Question Period	-0.336				
	(0.390)				
$I(PM_{2.5} \geq 15)$*Question Period	0.182				
	(0.231)				
Member of Government		0.262			
		(0.189)			
$I(PM_{2.5} \geq 15)$*Government		-0.527			
		(0.271)			
$I(PM_{2.5} \geq 15)$*I(Birth Year $>$ 1969)			-0.041		
			(0.326)		
$I(PM_{2.5} \geq 15)$*I(Birth Year $>$ 1959)				0.096	
				(0.302)	
$I(PM_{2.5} \geq 15)$*I(Birth Year $>$ 1949)					0.333
					(0.387)
Number of MPs	488	488	297	297	297
Observations	119,225	119,225	71,408	71,408	71,408

Values in parentheses are standard errors clustered on individual MPs.

All models contain fixed effects for day of week, month, and MP, as well as weather controls.

A Additional Detail on Speech Codification

This appendix describes some additional details on the method used to collect speeches from the Canadian House of Commons.

Computer code written in the Python language processed all interventions made in the Canadian House of Commons between 2006 and 2011. The code counts the number of sentences in each speech, the number of words, the number of characters and the number of syllables per word. The syllable counting regime is as follows. First, each word is looked-up in Python's Natural Language (Carnegie Mellon) dictionary. If the word is in the dictionary, the dictionary pronunciation guide is used to estimate syllables. If the word is not contained in the dictionary, it is separated into vowel clusters. As an example, "turtle" has two vowel clusters – "u" and "e" – and therefore two syllables. Syllables in "turtle" would be counted properly. However, there are exceptions such as "delicious". In this case, the third vowel cluster is actually two syllables, but the algorithm estimates three syllables when there would actually be four ("dee" + "lish" + "ee" + "us"). A cursory inspection led us to believe that the number of exceptions – i.e., words not contained in a dictionary *and* those not correctly coded according to vowel clusters – is small and would require significant manual processing to rectify. Therefore, we maintain the count created by the algorithm. These sentence, word, character and syllable count statistics are sufficient to construct various measures of text complexity. Simple measures, for example, are the number of words per sentence or the number of syllables per word. Slightly more complex statistics, like those used in the main text and in Appendix B, include the Flesh-Kincaid grade level score. These readability scores combine the components of a speech into an index of speech complexity. The Python code will be included with the data.

Finally two additional dummy variables are created. First, an indicator distinguishes between a speech made in Question Period versus a committee meeting or alternative debate. Further, it is straightforward to determine whether the speaker is a member of the Government or Opposition.

B Other Speech Complexity Indexes

Our primary econometric specifications use the Flesch-Kincaid grade level index as discussed in section 2.2.1. However, Table 3 includes results using four alternative readability indices. The formulas for these weighting schemes are as follows.

Coleman-Liau index

$$CLI = 0.0588L - 0.296S - 15.8$$

where L is the average number of letters per 100 words and S is the average number of sentences per 100 words.

Flesch Reading Ease Score

$$FRE = 206.835 - 1.015ASL - 84.6ASW$$

where ASL is average sentence length (number of words divided by number of sentences) and ASW is average word length in syllables (number of syllables divided by number of words).

Automated readability index

$$ARI = 4.71\left(\frac{\text{characters}}{\text{words}}\right) + 0.5\left(\frac{\text{words}}{\text{sentences}}\right) - 21.43$$

where characters is the number of letters and numbers, words is the number of spaces, and sentences is the number of sentences.

SMOG index

$$SMOG = 1.043 * \sqrt{\left(\text{Number of Polysyllables} \cdot \frac{30}{\text{Number of Sentences}}\right)} - 3.1291$$

Polysyllables are words with three or more syllables. The standard SMOG index only uses texts with 30 or more sentences, a restriction that would yield zero observations in our dataset.

C Additional Results

This section provides additional results as described in the main text.

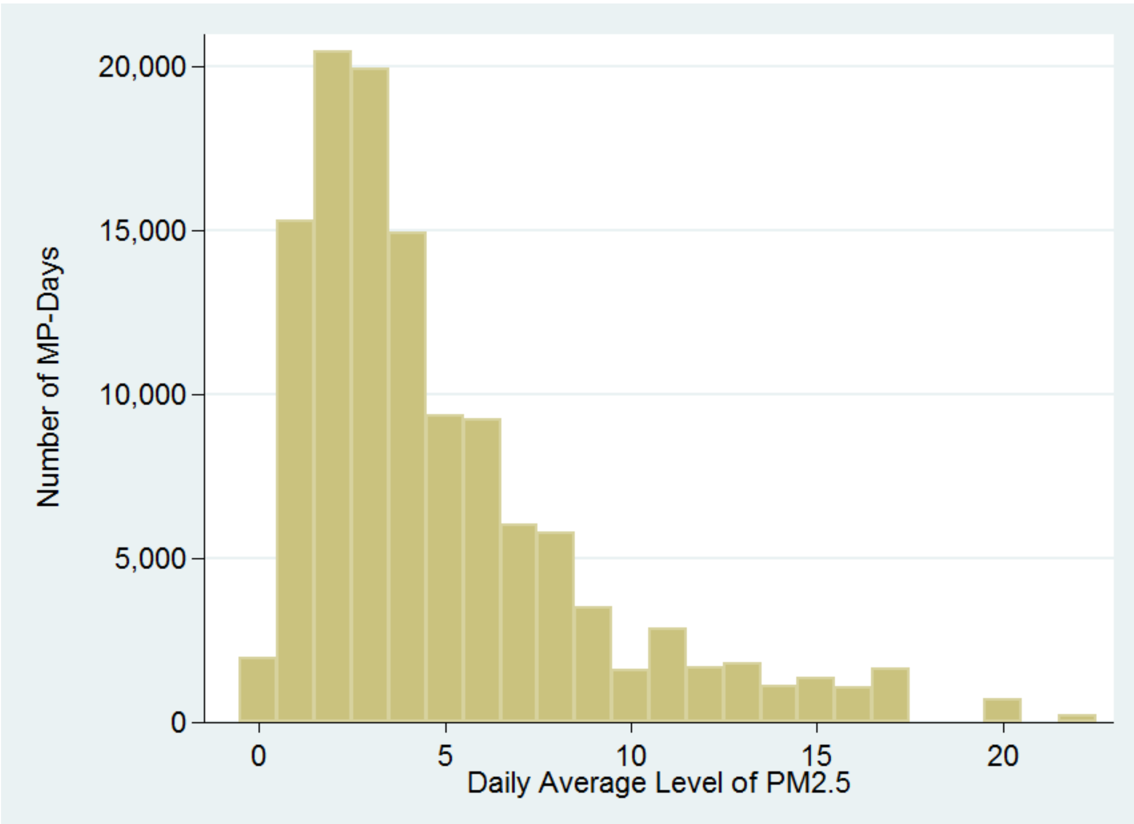

Figure 2: Distribution of Canadian MP-Days at Distinct Daily Average $PM_{2.5}$ Levels

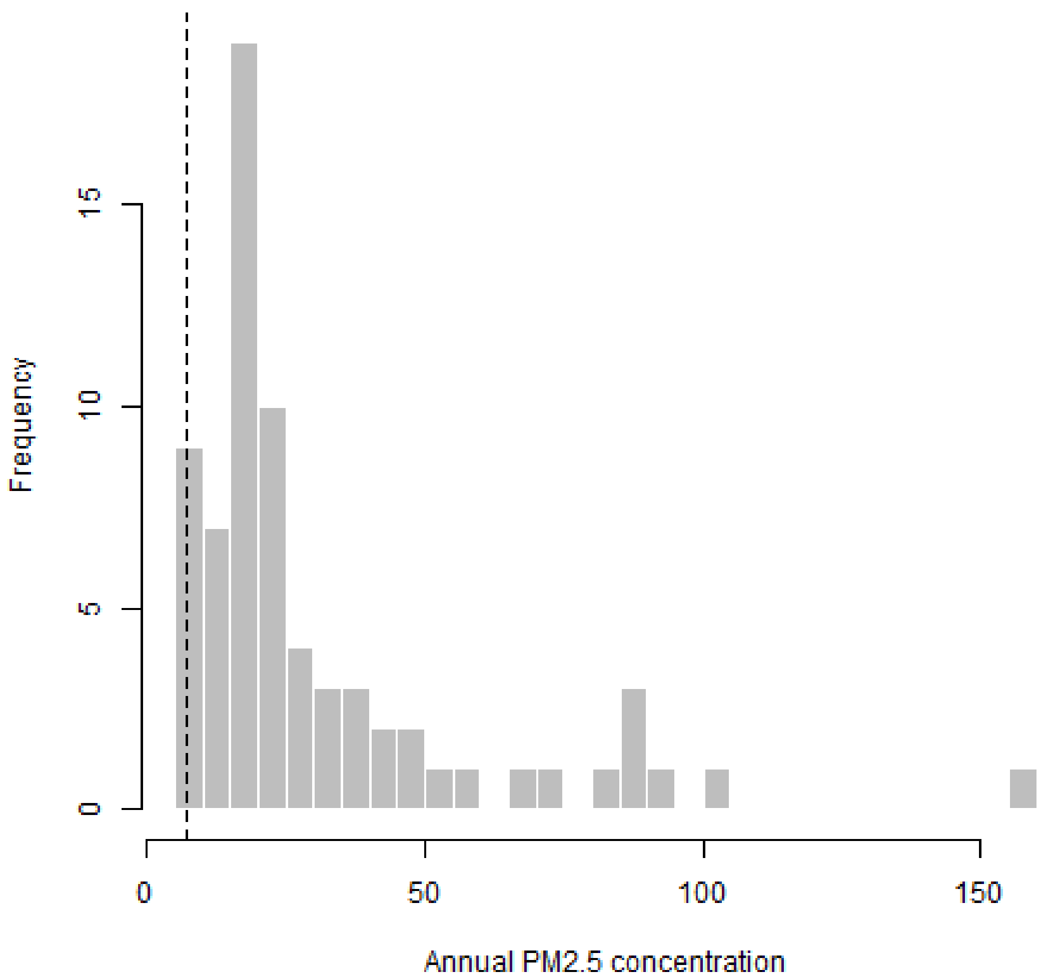

Figure 3: Distribution of National Capitals by Average $PM_{2.5}$ Levels

This figure plots the mean $PM_{2.5}$ concentrations for international capital cities. Ottawa is marked with the dashed line. Data is from the World Health Organization Ambient Air Polllution database, available at `www.who.int`.

Table 6: Effect of Elevated Particulate Matter Pollution Concentrations on the Speech Complexity of Canadian Members of Parliament, Alternative Specifications for Pollution Exposure

	Flesch-Kincaid Index			log(*Flesch-Kincaid Index*)		
Panel A: Linear, Continuous Pollution Exposure						
$PM_{2.5}$	-0.017	-0.010	-0.009	-0.001	-0.001	-0.001
	(0.007)	(0.009)	(0.009)	(0.001)	(0.001)	(0.001)
Panel B: Flexibly Binned Pollution Exposure						
$I(5 \leq PM_{2.5} < 10)$	-0.112	-0.079	-0.079	-0.013	-0.011	-0.011
	(0.055)	(0.061)	(0.061)	(0.005)	(0.005)	(0.005)
$I(10 \leq PM_{2.5} < 15)$	0.021	0.062	0.064	0.009	0.008	0.008
	(0.095)	(0.105)	(0.105)	(0.009)	(0.010)	(0.010)
$I(PM_{2.5} \geq 15)$	-0.432	-0.333	-0.324	-0.038	-0.036	-0.035
	(0.152)	(0.172)	(0.172)	(0.018)	(0.021)	(0.021)
Weather controls		✓	✓		✓	✓
Day-of-week fixed effects		✓	✓		✓	✓
Month fixed effects		✓	✓		✓	✓
MP fixed effects			✓			✓
Number of MPs	488	488	488	480	480	480
Observations	119,225	119,225	119,225	110,913	110,913	110,913

Values in parentheses are standard errors clustered on individual MPs.

This table replicates the results from Table 2 using alternative specifications of the econometric model. The Panel A includes pollution exposure as a linear and continuous covariate. The Panel B includes dummy variables for two additional pollution exposure bins. The results corroborate the conclusions in the main text and Figure 1.

Table 7: Weather Coefficients for our Preferred Specification: Column (3) from Table 2

Temperature (°C)	0.004
	(0.004)
Temperature^2 (°C) [x10^2]	0.008
	(0.028)
Rain (mm)	-0.007
	(0.004)
Rain^2 (mm) [x10^2]	0.008
	(0.006)

Values in parentheses are standard errors clustered by MP.

Temperature refers to the mean daily temperature. Rain is the cumulative daily rainfall. The coefficients and standard errors for both temperature squared and rain squared are scaled by 100.

D Implications of Bias due to Measurement Error

This section describes the reasoning for our claims about the prospective biases in our coefficients due to measurement error. The formalization is based on Bound et al. (1994). Let the measured speech complexity index and pollution levels, respectively, be written as:

$$y = y^* + \nu \tag{6}$$

$$p = p^* + u \tag{7}$$

where y is the speech complexity index (i.e., the Flesch-Kincaid index) and p is a measure of particulate matter (i.e., $PM_{2.5}$). An * indicates the true value of the data – i.e., the spoken level of complexity and individual level of pollution exposure. ν and u are error terms that are (potentially) correlated with these true values. Our preferred regressions are in levels with a full suite of time and individual fixed effects included. These parameters remove any time invariant individual measurement error (e.g., if an MP repeatedly uses a word incorrectly) or variation that is common to all MPs (e.g., if pollution is systematically higher on, say, Mondays). For current purposes, we ignore these effects and focus on the two aforementioned sources measurement error arising from the data-interpretation mismatch. For the dependent variable, this is Hansard editing and, for the independent variable, inexact assignment of pollution exposure.

We start with the error in the Flesch-Kincaid index and make several observations. First, editors do not transcribe texts on the same day on which the words are spoken. While we do not have information on specific transcription dates, we believe it is reasonable to treat ν as independent of pollution, p and u. Next, our research hypothesis is that MPs will be affected by pollution. This phenomenon may manifest itself in several ways. MPs may stumble or have a greater propensity to use verbal ticks. These "ums" and "ahs" are then systematically edited out of the recorded text in non-classical fashion. This means that the level of editing applied to a specific speech is correlated with the true level of speech complexity. Hence we rewrite ν

from (6) as:

$$\nu = \delta y^* + \nu^* \tag{8}$$

where ν^* is uncorrelated with the dependent (and independent) variable and δ is the coefficient from a hypothetical regression of ν on the true speech quality index, y^*. Based on what we know about the editing process, we expect that $\delta \leq 0$. This is because when MPs increase the frequency of "ums" and "ahs" editors will be active (i.e., short, single syllable words are deleted from the official text) and y^* will be low. Recording a $y > y^*$ implies that $\nu > 0$ for low values of y^*, so regressing ν on y^* gives a coefficient, δ, which is less than zero.

We next turn to pollution assignment. Given our design, we maintain that assignment of pollution exposure is conditionally independent of potential outcomes – politicians are making speeches for citizens who live across the country and their statements are formally documented within the official record (the database that we exploit). Still, there may be error in the measurement of the pollution assigned to specific MPs. Prior to making a speech in the House of Commons, an MP may have travelled to a heavily polluted location or may have time-varying health issues that make her more susceptible to ambient concentrations on a particular day. Pollution levels vary throughout the day, so averages may over- or under-state true exposure. Moreover, we focus on contemporaneous pollution and lagged exposure may matter. Overall however we treat the error in pollution assignment, the independent variable of interest, as classical errors-in-variables – i.e., u is uncorrelated with p^*. This errors-in-variables specification implies attenuation bias that is proportional to the ratio of the variance of u to the variance of the measured p. The magnitude of this bias is captured by the coefficient a hypothetical regression of u on p. Define λ as the estimated coefficient from this (hypothetical) regression. And as we are dealing with attenuation bias, we expect that $\lambda \leq 1$.

We now combine these two biases. Let the true parameter from a linear least squares regression of the Flesch-Kincaid index on pollution concentration equal β (i.e., this is the parameter that we would estimate without measurement error) and what we actually estimate be $\hat{\beta}$.

Using (8) and (7), we can write the bias in the estimated parameter as:[19]

$$\begin{aligned} \text{plim } \hat{\beta} - \beta &= -\lambda\beta + \delta\beta \\ &= -\beta\left(|\delta| + \lambda\right) \end{aligned} \tag{9}$$

where the second line follows from $\delta \leq 0$. The bias in (9) equals $-\beta\left(|\delta| + \lambda\right)$ and shows the attenuation arising from the biases in the dependent variable and pollution assignment. Error in the dependent variable leads to an downward (toward zero) bias of δ whereas λ reflects the conventional attenuation bias (also towards zero) of the standard errors-in-variables model. Both biases are then scaled by the true effect size, β.

[19] Using the following assumptions – $\text{cov}(p^*, \nu^*) = 0$ and $\text{cov}(p^*, \epsilon) = 0$ where ϵ is the conventional mean zero error term – the standard derivation of $\beta - \hat{\beta} = (p'p)^{-1}p'y$ – and defining λ as the standard errors-in-variables ratio of variances, this expression is derived as follows:

$$\begin{aligned} \text{plim } \hat{\beta} &= \text{plim } (p'p)^{-1}p'\hat{y} \\ &= \text{plim } (p^{*\prime}p^*)^{-1}p^{*\prime}(y^* + \epsilon - u\beta + \nu) \\ &= \beta + 0 + \text{plim } (p^{*\prime}p^*)^{-1}p^{*\prime}(-u\beta + \nu) \\ &= \beta + \text{plim } (p^{*\prime}p^*)^{-1}p^{*\prime}(-u\beta) + \text{plim } (p^{*\prime}p^*)^{-1}p^{*\prime}(\delta y^* + \nu^*) \\ &= \beta + \beta \text{ plim } (p^{*\prime}p^*)^{-1}p^{*\prime}(-u) + \delta \;\cdot \beta + \text{plim } (p^{*\prime}p^*)^{-1}p^{*\prime}(\nu^*) \\ &= \beta + \beta(-\lambda + \delta) \end{aligned}$$