What factors affect physicians' labour supply: Comparing structural discrete choice and reduced-form approaches[*a]

Guyonne Kalb

Melbourne Institute of Applied Economic and Social Research

The University of Melbourne and Institute for the Study of Labor (IZA)

Daniel Kuehnle

School of Business and Economics

FAU Universität Erlangen-Nürnberg

Anthony Scott

Melbourne Institute of Applied Economic and Social Research

The University of Melbourne

Terence Chai Cheng

School of Economics, The University of Adelaide

Sung-Hee Jeon

Statistics Canada

This version: August 25, 2017

[*] Note that all Appendix tables are available in the online Appendix.

[a] We are grateful for helpful comments and suggestions received by Richard Blundell, Jongsay Yong, an anonymous referee and participants at the 2014 European Association of Labour Economics meeting, the 2014 IZA European Summer School in Labor Economics, the Brown Bag Series at the Melbourne Institute, and the doctoral workshop at the University of Erlangen-Nürnberg. This work was supported by a National Health and Medical Research Council Health Services Research Grant (2007-2011) and a National Health and Medical Research Council Centre of Research Excellence in Medical Workforce Dynamics Grant (2012-2016) with additional support from the Department of Health (in 2008) and Health Workforce Australia (in 2013). None of the funding sources have had any involvement in the actual research or writing of the paper. The views in this paper are those of the authors alone. We thank the doctors who gave their valuable time to participate in the Medicine in Australia: Balancing Employment and Life (MABEL) survey, and the other members of the MABEL team for data cleaning and comments on drafts of this paper. The study was approved by the University of Melbourne Faculty of Business and Economics Human Ethics Advisory Group (Ref. 0709559) and the Monash University Standing Committee on Ethics in Research Involving Humans (Ref. CF07/1102 2007000291).

Abstract

Little is known about the response of physicians to changes in compensation: do increases in compensation increase or decrease labour supply? In this paper, we estimate wage elasticities for physicians. We apply both a structural discrete choice approach and a reduced-form approach to examine how these different approaches affect wage elasticities at the intensive margin. Using uniquely rich data collected from a large sample of General Practitioners (GPs) and specialists in Australia, we estimate three alternative utility specifications (quadratic, translog and box-cox utility functions) in the structural approach, as well as a reduced-form specification, separately for men and women. Australian data is particularly suited for this analysis due to a lack of regulation of physicians' fees leading to variation in earnings. All models predict small negative wage elasticities for male and female GPs and specialists passing several sensitivity checks. For this high-income and long-working-hours population, the translog and box-cox utility functions outperform the quadratic utility function. Simulating the effects of 5 and 10% wage increases at the intensive margin slightly reduces the full-time equivalent supply of male GPs, and to a lesser extent of male specialists and female GPs.

Keywords: labour supply, discrete choice model, wage elasticity, health workforce, MABEL survey.
JEL-codes: I11, J22, J44, J21.
Conflict of interest: no conflicts of interest are present for any of the authors.

1 Introduction

A critical task for health care policy is to ensure the long-term supply of medical services for an ageing population and an increasing demand for health care. To address the seemingly persistent disequilibrium in physician labour markets (World Health Organisation, 2013), national policy responses have largely focused on changing the number of medical graduates (McPake, Scott, and Ijeoma, 2014), the geographic distribution of physicians and the use of financial incentives, particularly pay-for-performance systems, to improve access to health care.

Though remuneration and financial incentives for physicians are increasingly focused on rewarding quality and performance, changes in earnings per hour from any payment method are also likely to influence the number of hours worked. For example, the introduction of pay-for-performance can increase earnings per hour and quality of care for patients, and through increased hourly earnings it can increase or decrease hours worked depending on whether the labour supply curve is backward bending. In turn, changes in hours worked influences access to care for patients as hours worked influences how many patients can be seen by a physician. Thus, if hours worked fall as earnings per hour increase, this could potentially reduce access, including equity of access, which is one of the key objectives in many health systems.

Despite a vast general labour supply literature (see e.g., Blundell, MaCurdy, and Meghir, 2007) and shortages of doctors in many developed countries, the labour supply of doctors has received surprisingly little attention; evidence exists only for two (very different) countries, the US and Norway. Most studies that examine doctors' labour supply apply a reduced-form approach specifying a log-linear relationship between hours worked and the wage rate (Sloan, 1975; Rizzo and Blumenthal, 1994; Showalter and Thurston, 1997; Thornton, 1998; Ikenwilo and Scott, 2007). However, this popular specification linearises the equation in log wages at the observed labour supply point, imposes a constant wage elasticity for all doctors and ignores potential heterogeneities. Imposing a constant wage elasticity is particularly problematic because wage elasticities should decline at higher hours of work due to an increased marginal utility of leisure relative to utility of income.

We provide evidence on the determinants of doctors' labour supply by using very rich data from a unique Australian study of doctors, “Medicine in Australia: Balancing Employment and Life” (MABEL). We estimate structural and reduced-form models, and compare the estimated wage elasticity from both types of models for doctors with different characteristics. Australian data are particularly suited to estimate physicians’ wage elasticities due to a lack of regulation of fees charged by physicians, which provides the required variation in earnings for estimation.

We contribute to the literature in four ways. First, we estimate a structural discrete choice labour supply model which has gained increasing popularity in the general labour economics literature. A few studies, e.g., Cheng, Kalb, and Scott (2017), Andreassen, Di Tommaso, and Strøm (2013), and Sæther (2005), apply a structural labour supply model to physicians but mainly examine choices between different types of jobs (e.g., public versus private). The discrete choice approach offers a number of advantages compared to the continuous approach, including the flexibility of the functional form, the relative ease of incorporating complex non-linear tax and transfer systems, the broader range of utility functions that can be used, and no need to impose quasi-concavity conditions a priori. The discretisation of working hours may be particularly appropriate for General Practitioners (GPs) who often have a limited number of working hours to choose from due to institutional constraints affecting their labour supply (for general populations and doctors, respectively, see: Van Soest, Woittiez, and Kapteyn, 1990; Sæther, 2005).

Second, we explore heterogeneous responses by providing a detailed analysis for different subgroups of doctors. Whereas previous studies have relied on small samples and estimated models for male and female doctors combined, or models for male doctors only, our larger sample allows for separate models by gender and doctor type. Given an increasing proportion of female doctors, it is important to estimate separate models by gender. Third, we explore the sensitivity of the models’ implications to the utility function specification and to using a different number of discrete labour supply choices. This is relevant to the broader labour supply literature as well, particularly in relation to high-income individuals. Finally, we use different specifications of the structural discrete choice approach to simulate labour supply responses at the intensive margin in response to wage increases of 5 and 10% taking into account the non-linearity of the tax schedule.

The paper proceeds as follows. Section 2 presents a brief literature review on physicians' labour supply and summarises the reported wage elasticities. This is followed by a brief description of the Australian institutional setting in Section 3. Section 4 outlines the two types of labour supply models and the associated estimation approaches. In Section 5 we describe the data and discuss some descriptive statistics. Section 6 presents the results, including a policy simulation. The paper concludes with a discussion of the implications in Section 7.

2 Literature review

The earliest studies on the determinants of doctors' labour supply, e.g. Feldstein (1970), Fuchs and Kramer (1972), Brown and Lapan (1972), run OLS regressions of the quantity of services provided by a GP on different control variables and a fee measure. These US studies generally find small negative fee elasticities that are measured imprecisely due to small sample sizes. Sloan (1975) uses US census data from 1960 and 1970 and finds small positive wage elasticities (< 0.1) on average as well as evidence for a backward-bending labour supply curve for a minority of doctors at the top of the income distribution.

Rizzo and Blumenthal (1994) model labour supply and the wage rate jointly based on a sample of young self-employed physicians from the 1987 Practice Patterns of Young Physicians Survey. The study estimates the model for men and women combined, and finds a positive wage elasticity of 0.23. Showalter and Thurston (1997) study the effect of changes in state marginal tax rates on labour supply and find significant positive wage elasticities for self-employed physicians (0.33), but small (0.10) and insignificant wage elasticities for doctors on wages or salaries. Using the same data Thornton (1998) estimates very small positive wage elasticities (0.06) for male, self-employed, solo-practice physicians and concludes that reductions in medical fees are unlikely to decrease the supply of medical services.

For Norway, Baltagi et al. (2005) use administrative data from 1993 to 1997 for male hospital physicians and apply different estimators to their labour supply model, exploiting a natural experiment in which some doctors received a 15% wage increase while others did not. Estimating the labour supply model by GMM, they find significant positive wage elasticities of around 0.3.

The above studies all use a reduced-form approach. Only a small number of studies use a structural discrete-choice approach. Using administrative data for Norwegian residents in 1995 and 1997, Sæther (2005) estimates a structural discrete choice labour supply model for doctors aged 28-66, both employed and self-employed. He finds wage elasticities for hospital physicians ranging broadly from 0.1 to 0.2. He also shows that private sector wage increases lead to stronger changes in hours worked in the relevant sector than public sector wage increases.

Most recently, Andreassen et al. (2013) use Norwegian administrative data from 1996-2000 to estimate a labour supply model that allows doctors to choose between 10 different job packages which are derived from a combination of attributes: part- or full-time work, hospital or primary care, public or private sector, with 'working in other sectors' and 'not working' representing the 9th and 10th package. The study focuses on all employed married physicians and finds an average wage elasticity of 0.04.

3 Institutional context

Physicians in Australia are paid in two main ways. First, medical specialists can be employed by public hospitals on a salary, or a salary with rights to practice. This enables them to either see private patients in public hospitals where the hospital receives the revenue and the physician may receive a salary bonus or in-kind benefit (e.g. attendance at conferences). Rights to private practice means they can also see private patients in private hospitals or in their own consulting rooms. Unlike the US, medical specialists in private practice do not organise into large medical groups, and largely practise on their own whilst sharing administrative support. All GPs work in private practice in small medical groups with nurses and sometimes allied health. GPs are gatekeepers to specialist care.

Second, both GPs and specialists who work in private practice can charge fees to patients, with no regulation as to the level of fees. Medicare is Australia's tax-financed universal health scheme, which provides a fixed subsidy to patients for each visit, procedure or test performed by GPs and specialists in private practice. The fixed subsidy is determined by the Medicare Benefits Schedule. If physicians charge above the subsidy, then patients face a co-payment which varies across physicians. This co-payment cannot be covered by private health insurance.

Eighty percent of GP visits are charged at the level of the Medicare subsidy, so patients do not face a co-payment. This is known as bulk-billing, and subsidies are higher for GP consultations for patients under 16 and patients with concession cards, providing incentives to bulk-bill these groups of patients. Only around 25% of specialist services are bulk-billed. There are additional incentive schemes including the Practice Incentive Program, to encourage GPs to provide certain services to patients, such as after hours care, chronic disease management, immunisations, cervical screening, and payments for working in non-metropolitan areas. Patients of GPs do not need to enrol or register with their GP.

Between 1997 and 2013, average working hours of Australian doctors have fallen from 48.0 to 42.8 hours. A number of factors contribute to the observed changes in working hours. First, the share of female doctors has increased substantially as women account for only 24% of doctors aged 55 and over, compared with 52.5% of doctors aged 35 and under. These changes in gender composition affect labour supply, since female doctors work 38.8 hours per week on average compared to 45.4 hours for males. Second, at the same time men have also reduced their working hours, and relatively more so than women over time.[1] Third, the proportion of older doctors has increased over the past ten years which resulted in a significant drop in average working hours as older doctors tend to work fewer hours. Fourth, Markwell and Wainer (2009) and Shrestha and Joyce (2011) document the changing work/life balance expectations amongst doctors. In addition, studies of retirement intentions suggest that one third of GPs plan to retire before age 65 (Brett et al., 2009).

4 Methods

4.1 A structural labour supply model

Our central analyses use a structural model of individual labour supply to estimate preference parameters and elasticities with respect to income and wages. We treat labour supply as a discrete choice problem rather than a continuous choice (e.g. Van Soest, 1995).

As in standard labour supply models, we assume that doctors choose a combination of

[1] All figures were obtained from the Australian Institute of Health and Welfare (AIHW, 2014, 2015).

hours worked and household net income that maximises their utility. We follow Löffler, Peichl, and Siegloch (2014) and compare three different utility functions: the quadratic, translog, and Box-Cox. The quadratic specification (e.g. Keane and Moffitt, 1998) is quite flexible, without imposing too many restrictions a priori, as individual leisure and consumption can be either substitutes or complements.[2] Furthermore, unlike other utility functions, the quadratic utility function takes working hours rather than leisure as its arguments and therefore does not require choosing an arbitrary value for the total endowment of time.

Alternatively, we estimate a translog utility function (e.g. Van Soest, 1995) which has the same advantages as the quadratic utility function except that it cannot be directly expressed in working hours but needs to be in terms of leisure. To compute leisure time we choose 168 as the total amount of time available per week. Finally, to ensure that the results are not driven by the choice of utility function, we also estimate a Box-Cox utility function (e.g. Sæther, 2005).[3] We again choose 168 hours per week as the total endowment of time. For all three utility specifications, we check post-estimation whether utility is quasi-concave at the observed labour supply point and thus consistent with economic theory (Varian, 1992, pp. 96-97).

We assume that each doctor i can choose between j alternatives from a limited set of m combinations of income and working hours, $\{(y_{ij}, h_{ij}); j = 1, 2, ..., m\}$, where y_{ij} is the household's net income, which includes other household income such as a partner's income or the doctor's non-labour income, associated with the doctor's working hours h_{ij}. We specify the three utility functions as follows.

The quadratic utility function:

$$U_{ij} = \beta_1 y_{ij} + \beta_2 y_{ij}^2 + \beta_3 h_{ij} + \beta_4 h_{ij}^2 + \beta_5 h_{ij} y_{ij} + \epsilon_{ij} \tag{1}$$

The translog utility function:

$$U_{ij} = \beta_1 lny_{ij} + \beta_2 (lny_{ij})^2 + \beta_3 ln(168 - h_{ij}) + \beta_4 (ln(168 - h_{ij}))^2 + \beta_5 ln(168 - h_{ij}) lny_{ij} + \epsilon_{ij} \tag{2}$$

[2] Van Soest et al. (2002) show that utility functions including fifth-order polynomials yield almost identical wage elasticities compared with models using second-order polynomials.

[3] Our Box-Cox utility specification includes an interaction term between income and hours worked that is absent in Sæther, and which turns out to matter substantially for the implied wage elasticities.

The Box-Cox utility function:

$$U_{ij} = \beta_1 \frac{y_{ij}^{\beta_2} - 1}{\beta_2} + \beta_3 \frac{((168 - h_{ij})/168)^{\beta_4} 1}{\beta_4} + \beta_5 \frac{y_{ij}^{\beta_2} - 1}{\beta_2} \frac{((168 - h_{ij})/168)^{\beta_4} 1}{\beta_4} + \epsilon_{ij} \quad (3)$$

In all three specifications, we assume that the random error ϵ_{ij} follows a type I Extreme Value distribution and estimate the parameters as a multinomial logit model by maximum likelihood. Furthermore, we always allow the vector of preference parameters β_1 and β_3 to differ by some individual characteristics, e.g., the number of children, the doctor's age, and health status.

Given the above models and assuming that individuals choose the alternative that leads to the highest utility, the probability that individual i chooses alternative j (from the m alternatives) is:

$$Pr(U_{ij} > U_{ik}, k \neq j) = \frac{exp(U_{ij})}{\sum_{k=1}^{m} exp(U_{ik})} \quad (4)$$

To estimate these probabilities we need to determine the utility level and thus the household net income associated with each choice j. To generate household net income, we first compute gross hourly wages either directly from observed information or by using wage regressions. Using gross hourly wages, we calculate gross labour income associated with each choice of working hours. We then add nonlabour income and the spouse's gross income to generate gross household income. Finally, we apply the Australian tax and transfer system to generate the amount of net household income associated with each level of working hours.[4]

Consistent with the previous literature (e.g. Sæther, 2005; Andreassen, Di Tommaso, and Strøm, 2013), we only vary policy parameters that affect the doctors and do not model their partners' labour supply choices which remain exogenous in our framework. Ideally, we would like to jointly model the labour supply of both spouses for couple families. Unfortunately, the data available does not provide information on partners' working hours.[5]

For our analysis, we choose discrete labour supply points that cover the observed labour supply as well as possible. Hence, our main model offers ten different choices of working

[4]We include individual income tax payments and income tax rebates, as well as Government payments to families with children.

[5]The lack of information on the partner's income may be less of an issue for men than for women. Our results which show that the partner's employment status affects labour supply by women but not men confirm this.

hours: 16, 20, 30, 40, 45, 50, 55, 60, 65 or 70 hours per week.[6] We also examine the sensitivity of results to choosing a smaller and larger number of labour supply points: five (allowing 20, 40, 50, 60 or 70 hours per week) and thirteen (allowing 8, 16, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 or 70 hours per week).

4.2 A reduced-form labour supply model

Starting from the same economic framework of utility maximisation (see Stern, 1986), and a few simplifying assumptions and approximations (such as linearising wages at the observed labour supply point), we derive a reduced-form static labour supply model as in equation 5:

$$ln(H_i) = \alpha_1 ln(W_i) + \alpha_2 ln(Y_i) + \mathbf{X}'\beta + \epsilon_i \tag{5}$$

where the natural logarithm of hours worked (H_i) is regressed on (the log of) the gross wage rate (W_i), (the log of) gross other non-labour income (Y_i), and a range of individual characteristics $\mathbf{X}$, e.g., the age of the doctor, number of children, age of the children. The parameter α_1 yields the uncompensated substitution elasticity (Blundell and MaCurdy, 1999, p.1599).

This reduced-form approach imposes a number of restrictive assumptions that the structural discrete choice model does not require. First, the model estimates a constant wage elasticity α_1. The linear specification is fairly restrictive as the wage elasticity may vary over the hours distribution or depend on non-labour income or other demographic characteristics. Second, the reduced-form model also assumes quasi-homothetic preferences which empirical studies on consumer behaviour (Blundell and MaCurdy, 1999, p.1687) typically reject. Third, the reduced-form specification includes the gross wage without allowing for non-linearity of the wage after applying the rules of the tax and transfer system. In effect, the equation is linearised in the wage at the observed labour supply point.

Despite these shortcomings, the model nevertheless provides an interesting benchmark against which to compare the average wage elasticity derived from the structural model. In addition, it allows for a comparison to the literature using the reduced-form approach.

[6] The corresponding hours intervals are: [0 -18); [18 -25); [25 -35); [35 -42.5); [42.5-47.5); [47.5-52.5); [52.5-57.5); [57.5-62.5); [62.5-67.5); [67.5-80).

4.3 The wage equation

Since we compute hourly wages by dividing total weekly income by the weekly hours of work (see Section 5.2), wages are likely to be endogenous and subject to measurement error transmitted through measurement error in weekly hours of work. By definition, we introduce spurious negative correlation between observed hourly wages and weekly hours of work whenever weekly hours of work are observed with error. Borjas (1980) explains how this relationship between observed hourly wages and hours worked, which is present in most labour supply studies, will bias wage elasticities downwards whenever measurement error is present in the hours of work variable.

To address the endogeneity of wages, we pursue three main strategies. First, we follow the literature (including Borjas, 1980) in using imputed wages instead of observed wage rates as our preferred specification. We estimate wage regressions separately by doctor type and gender following Cheng et al. (2012) including several variables that are not included in the preference parameters β_1 and β_3 of the utility functions in equations (1), (2) or (3), or in $\mathbf{X}$ in equation (5) as exclusion restrictions. These include attending medical school in Australia, number of postgraduate qualifications, type of postgraduate qualification, temporary visa holder, work experience, state of residence, remoteness indicators, local median house price, practice size (only for GPs), main speciality type (for specialists only), and percentage of time in clinical work (see Table A.I). To compute gross income at each labour supply point based on the parameter estimates from the wage equation, we predict gross hourly wage rates, multiply these with the hours worked at the relevant labour supply point, and compute other (non-labour and partner's) income in the same way as for the observed wage approach.

Our main approach involves the estimation of a wage equation by ordinary least squares using the first wave of data:

$$w_i = \mathbf{x}_i'\beta + \varepsilon_i \tag{6}$$

where w_i captures gross hourly earnings for individual i, the vector $\mathbf{x}_i$ consists of a number of explanatory variables (including the variables listed as exclusion restrictions above), and ε_i is a normally distributed error term. In the reduced-form specification of labour supply we can test for relevance and validity of the instruments through a first-stage F-test and the Sargan over-identification test respectively.

Second, to address the concern that the exclusion restrictions of the main wage imputations may not hold, we estimate an alternative version of the wage equation where we use the same structural and reduced form models as before, but now base the labour supply estimations on the third wave of the MABEL data. In this specification, we use lagged observed wages from waves 1 and 2 as instruments for observed wages in wave 3 to avoid reverse causality as suggested by Reed (2015). This approach also deals with unobserved heterogeneity given instrument validity which we can assess through an over-identification test.

Estimating wage equations for relatively homogeneous groups like medical doctors/specialists can be difficult. For example, there is unlikely to be much variation in education level, which traditionally explains a substantial proportion of wage variation between individuals. Therefore, we check the sensitivity of results to a third alternative specification of the wage equation based on more data, utilising the first four waves instead of just one wave. This uses additional information without relying on time periods that are too distant from wave 1:

$$w_{it} = \mathbf{x}'_{1it}\beta_1 + \mathbf{x}'_{2i}\beta_2 + c_i + \varepsilon_{it} \tag{7}$$

where w_{it} captures hourly earnings for individual i at time t, c_i denotes the unobserved individual-specific effects (or heterogeneity), and the vectors $\mathbf{x}_{1it}$ and $\mathbf{x}_{2i}$ consist of time-varying and time-invariant regressors, respectively.

Wages for the same doctor in different years are correlated, so it is necessary to control for this correlation in our estimation. The β_1 parameters of the model can be consistently estimated using conventional estimators such as fixed-effects or first-differencing estimators. These approaches, however, are not practical in our application as they do not permit estimating coefficients on time-invariant covariates (e.g. medical speciality, medical school in Australia), which are important predictors of doctors' earnings.

An approach that accommodates both individual heterogeneity and time-constant covariates is the correlated random effects (CRE) model, originally proposed by Mundlak (1978), and extended by Chamberlain (1982). Decomposing the individual heterogeneity term $c_i = \psi + \bar{\mathbf{x}}_i'\xi + a_i$ where $\bar{\mathbf{x}}_i = T^{-1}\Sigma_{t=1}^{T}\mathbf{x}_{1it}$, and ψ is some constant, w_{it} can be written as

$$w_{it} = \psi + \mathbf{x}_{1it}'\alpha_1 + \mathbf{x}_{2i}'\alpha_2 + \bar{\mathbf{x}}_i'\xi_1 + a_{1i} + \varepsilon_{it} \tag{8}$$

where it is usually assumed that $E(a_i|\mathbf{x}) = 0$ and $E(\varepsilon_{it}|\mathbf{x}) = 0$. The parameters of the model may be estimated using the pooled ordinary least squares estimator, which produces fixed effects estimates of the time-varying coefficients, as well as estimates of the time-invariant variables. For unbalanced panels, which applies in our context, $\bar{\mathbf{x}}_i$ is calculated as the time-average of $\mathbf{x}_{it}$ over the number of time periods that are observed for each individual i (Wooldridge, 2010). We report these additional robustness tests alongside our main results in Section 6.

5 Data and summary statistics

5.1 MABEL survey

This paper uses a unique longitudinal survey of doctors, "Medicine in Australia: Balancing Employment and Life" (MABEL), which covers many topics related to labour supply, e.g. characteristics of the work setting, workload, income, geographic location, demographic characteristics, and family circumstances. Joyce et al. (2010) provide a detailed discussion of the study design and show that the cohort is nationally representative with respect to age, gender, geographic location and hours worked. In our main analysis, we use data from the first wave of the MABEL survey, conducted in 2008, on qualified GPs and specialists *working in clinical practice*. Selecting doctors working in clinical practice only means that we can examine labour supply responses at the intensive margin but cannot analyse the decision to work in clinical practice.[7] For additional robustness checks concerning wage endogeneity, we also use data from waves 2 to 4.

[7]However, given the high investment in human capital required to become a doctor, relatively few qualified doctors do not work in their profession. The AIHW (2010) reports that about 7% of all registered medical practitioners do not work in the medical workforce. This number includes non-GPs and non-specialists, for whom the non-participation rate may be higher than for GPs and specialists.

5.2 Construction of income variables

To construct the key argument in the utility function, net income at each labour supply point, we need information on i) the gross hourly wage earned in medical practice and ii) gross other household income. The MABEL survey collects information on gross or net income reported per fortnight or annually, and separately asks for income from the medical practice and for total household income.[8] If doctors provide weekly or fortnightly income figures, we assume that this income was the same over all weeks/fortnights worked to impute an annual income value. We divide annual medical income by annual hours worked in the medical practice to compute the gross hourly wage earned in medical practice. We either use observed gross hourly wages directly in the labour supply model, or as the dependent variable in the wage equations to predict hourly wage rates as described in Section 4.3.

We compute gross other household income by subtracting the income from medical practice from total household income. Other household income thus includes the doctors' income from other sources and, for cohabiting doctors, the partner's labour and non-labour income, or a mix of these sources.

Using the relevant tax and family support rules from 2008, we compute net income from gross income. Because of individual taxation and a lack of detail about the partner's income, we are required to make a few assumptions about the split of other household income. First, if the partner is working (either full- or part-time), we allocate all other income entirely to the doctor's partner. Second, if the partner is not working, then we split the other household income equally between both spouses assuming that the couple aims to maximise tax benefits (e.g., to use the tax-free income threshold).

To address the sensitivity of results with respect to measurement error in the partner's income or other household income, we also apply three alternative approaches (briefly described in Table A.II) to construct these two measures of income both when using observed wages and imputed wages. The estimated wage elasticities from these alternative approaches are very similar to those from the base case approach.[9] Thus, our results are robust to the different ap-

[8]Although the response rate for the financial variables is lower than for some of the other questions (Kuehnle et al., 2010), the majority of GPs and specialists (85.3%) provide either gross or net income.

[9]Results based on observed wages (circumventing the need to estimate three additional sets of wage equations)

proaches taken to compute the doctor's medical earnings, the division between the partner's earnings and other household income.

5.3 Summary statistics

We present descriptive statistics for our estimation sample on average hours worked by gender, doctor type, and age in Figure 1, together with the proportions of women in each age group. First, the proportion of women decreases with age and is largest amongst the younger cohorts. Second, men and women differ markedly in their labour supply over the life-cycle. For instance, women in their prime child-rearing ages (30-49) work the lowest average hours. Conversely, women aged 50-59 work the longest hours amongst women, which is likely to be due to children having grown up by this stage.

Figure 2 presents the distribution of observed working hours by gender and doctor type. It shows that women work fewer hours than men and, for GPs and specialists, women represent the majority of the part-time doctors (e.g. less than 40 hours). Second, specialists are more likely to work long hours than GPs. Furthermore, the observed distribution of hours worked suggests that a broad range of working hours is on offer to doctors, facilitating the supply of preferred hours without facing major demand side constraints. [10]

Table I contains the summary statistics for all variables used in the analysis and reveals several differences in socio-economic characteristics between the four groups of doctors. As expected, specialists earn more per hour than GPs, and in both groups women earn less per hour than men (also see Schurer et al., 2016). This pattern can be clearly observed in Figure 3 which also reveals an inverse-U relationship between wages over the life cycle of male doctors, where wages are highest for the age group 50-59. For women, wages vary much less with age, with the one exception of female GPs aged 70-79 which is, however, a small group of GPs.

are available in Table A.II (rows 2 to 4).

[10] Table A.II (panel E) and accompanying description show that the estimation results remain of the same order of magnitude (although mostly insignificant) after dropping individuals who are most likely to face demand side constraints following an approach by Ribeiro (2001).

6 Results

6.1 Performance of alternative utility function specifications

We begin our analysis with a comparison of the goodness of fit measures for the three different utility functions. In contrast to other studies (e.g. Van Soest, 1995; Hanel, Kalb, and Scott, 2014), the quadratic utility function violates the quasi-concavity conditions in the majority of cases for our sample of highly-paid doctors working long hours. Table II shows that the utility function is quasi-concave for only 17% of female GPs, 57% of female specialists, 15% of male GPs, and 23% of male specialists. The other two utility functions perform much better in this regard. When using the translog specification, the quasi-concavity conditions are fulfilled for the vast majority of observations. The results for the Box-Cox utility are quite similar, but we were not able to estimate a Box-Cox model with the interaction term between income and hours worked included for female GPs.[11] Given the quadratic specifications' failure to fulfill the quasi-concavity conditions for a large number of observations and the difficulties in getting the Box-Cox models to converge, we use the translog utility function as the benchmark specification in this paper.[12]

With regard to the other goodness-of-fit measures, there is relatively little difference between the specifications, or the preferred specification is a different one for the different doctor groups. The average predicted hours are nearly all the same in the different utility specifications, with predicted hours being slightly higher than observed hours in all groups and for all specifications. The proportion that is correctly predicted is quite similar for the different specifications (with the highest proportion being for a different specification for the different doctor groups). The proportion that is correctly predicted is higher for women than for men, and is higher for GPs than for specialists across all specifications. Based on the Akaike and Bayesian Information Criteria, a different specification is preferred for each of the doctor groups.

[11] Compared to the quadratic and translog specifications, convergence of the model was much more difficult to achieve for the Box-Cox specification for all doctor types.

[12] Since the Box-Cox model for female GPs did not converge, we report results from the Box-Cox model without interaction terms between income and hours worked for this group. However, note that results for the other groups (Table IV, panel B) show that excluding the interaction reduces the size of the predicted wage elasticity.

6.2 Estimated marginal effects on labour supply

This section discusses the results from the structural labour supply model with 10 discrete hours points based on the translog utility function (for coefficients, see Table A.III). We present simulated marginal effects (expressed in absolute hour changes) and their 95% confidence intervals in Table III.[13,14]

Table III reveals interesting similarities and differences between the four doctor groups. As expected, young children reduce working hours for all groups; this reduction is largest for female GPs (who are expected to work on average 11.52 hours less if a child aged 0-4 is present compared to having no children under 16), and then female specialists and male GPs. For women, the effect of the total number of children, although insignificant, compounds the negative effect of the youngest child, while for men the effect of family size is positive and significant, thus making the combined effect of the child variables ambiguous and dependent on the age and number of children. Our results are consistent with the findings by Wang and Sweetman (2013).

Reflecting the observed decline in working hours across the age distribution, increasing age by one year decreases labour supply for all doctor types, except for female specialists. Health status appears somewhat important for GPs but not for specialists. Having good health instead of very good or excellent health increases the expected hours of work, especially for female GPs (note that we cannot exclude the possibility of reverse causation). Few doctors have poor or fair health.

The marginal effects of having a partner reveal some interesting patterns. If the partner is not employed, female doctors tend to work more hours than single female doctors, while it makes no difference to male partnered doctors compared to single male doctors. Men generally seem non-responsive to their partner's working status. If the partner is in full-time employment, female specialists and GPs work slightly fewer hours compared to single women.

Finally, self-employed doctors and GPs working in remote areas of Australia work more

[13]We compute the 95% confidence intervals using the percentile method within a parametric bootstrap approach (Cameron and Trivedi, 2005).

[14]In further sensitivity checks, we show that the results are robust to changes in the utility function (Table A.IV), estimation approach (Table A.V) and the wage definition (Table A.VI).

hours than other doctors.

6.3 Wage elasticity

In this section we use the estimated parameters for a range of different specifications to predict individual doctors' responses to a 1% increase in individual wages interpreting these as the wage elasticities.[15] Table IV reports average elasticities for each of the specifications.[16]

A number of important points stand out. First, we observe negative average wage elasticities for male and female doctors, GPs and specialists, indicating that the estimated working hours of many Australian doctors correspond to the backward bending parts of their labour supply curves. The elasticities are modest in size and range in value between -0.06 and -0.23. The negative wage elasticities are mostly significant for both men and women, except for the estimates using imputed wages for female specialists.

Second, Panel A shows that the negative wage elasticities are not driven by the choice of the number of discrete labour supply points allowed in the specification of the discrete choice model. Five, ten or thirteen mid-points yield very similar results, except perhaps for female GPs for whom the model with 5 mid-points appears to introduce substantial measurement error by not covering the observed distribution of labour supply well enough.

Third, the estimated negative wage elasticities are quite robust on average against using observed or imputed wages. The point estimates are never significantly different from each other, although some of the estimates using imputed wages are not significantly different from zero due to the loss of precision.

Fourth, Panel B shows that the implied wage elasticities are very similar when using the quadratic utility function or the full Box-Cox utility function (with interaction terms between income and hours) as compared to using the translog function. This similarity is surprising given the poor performance of the quadratic utility function in terms of meeting the quasi-concavity conditions at the observed labour supply points. The largest difference is again observed for female GPs. The reduced Box-Cox specification, which omits the interaction between income

[15] For self-employed doctors who are not paid an hourly wage but are reimbursed for patients seen or services delivered, we can interpret this as a 1% increase in the reimbursement fee.

[16] A similar approach as described in footnote 13 is used to compute the 95% confidence intervals.

and hours, converges for all groups but produces substantially smaller, but still negative, wage elasticities.

Fifth, the table shows that structural and reduced-form approaches (Panel C) produce strikingly similar wage elasticities on average for each of the four subgroups, except for female GPs.[17] The similarity indicates that the constant wage elasticity estimated in the reduced-form approach is consistent with the average elasticity in the structural discrete choice approach.

In addition to these main results, we carry out several sensitivity checks, including: using lagged wages as instruments for current wages, using four waves of data to impute wages, accounting for wage prediction errors and working hours mismatches, using alternative wage definitions, and investigating differences by employment status. Table A.II in the Appendix shows that our main results are robust to these additional checks which also alleviate concerns that measurement error and wage endogeneity might bias our findings.[18]

The advantage of using a structural approach becomes clear when we present the variation in estimated wage elasticities of individual doctors from the baseline model graphically as in Figure 4. The figure clearly shows that wage increases lead to heterogeneous responses. These cannot be incorporated in the reduced-form model, but can be reflected through the structural model. While the probability mass is mostly to the left of zero, reflecting negative wage elasticities on average, a small proportion of doctors are estimated to have positive wage elasticities.

Finally, we want to reveal the potentially heterogeneous effects for sub-populations which health authorities could target specifically. Therefore, Figure 5 presents the estimated wage

[17]Given the large proportion of female GPs working part-time hours, there is likely to be more variation in wage elasticities between female GPs than within the other groups. As a result, the assumption of a constant wage elasticity across all female GPs (implied by the reduced-form approach) may be a more restrictive assumption for this group.

[18]In the context of the reduced-form approach, the validity of the instrumental variables that we use to predict wages in our main approach can be assessed through an over-identification test. While the instruments of our main specification do not pass the over-identification test, once we use a quadratic experience term in the wage equation, leave out "percentage of time in clinical work", and omit "practice size" (for GPs) and "specialty dummies" (for specialists), the test is passed. When re-estimating the wage elasticities in the structural models using this specification, we obtain results that are similar to our main results (results are available upon request). Thus, while our main specification does not pass the Sargan test, we reach the same conclusions regarding the negative wage elasticities when using a more parsimonious specification of the structural model or when using lagged wages as instruments, with both specifications passing the Sargan test in the corresponding reduced-form context. However, the parsimonious specification leaves out important variables which are particularly relevant for the more rigid reduced-form model. Borjas (1980) noted a similar sensitivity to inclusion/exclusion of specific variables in reduced-form labour supply models. Therefore, we prefer to keep these in our main model.

elasticities for a number of selected subgroups. Generally, the groupings in the figure do not clearly identify particular groups that would respond more strongly to wage increases than other groups. The subgroup analysis shows that the labour supply of specialists (and particularly male specialists) does not respond much to wage increases. As expected, male and female GPs working longer hours have on average slightly larger negative wage elasticities than doctors working fewer hours. The only subgroups that stand out slightly are female specialists with a non-employed partner who are the only group to respond positively on average, although insignificantly, to a 1% increase in wages. Doctors with a preschool child (except for male specialists) have a very small negative average wage elasticity close to zero, and are more likely to have a positive wage elasticity than other groups.

6.4 Policy simulations

Finally, we use the structural model to simulate doctors' labour supply responses to additional increases in the wage rate of 5% and 10%. We calculate the labour supply responses as the percentage change in hours per week and in terms of the absolute change in full-time equivalent (FTE) doctors. FTE is a meaningful measure of supply because it takes into account differences in hours worked among doctors. We calculate the FTE measure by multiplying the number of medical practitioners in the population by the average change in weekly hours worked, and dividing the result by the number of hours in a standard full-time working week.[19]

The simulation results are shown in Table V together with the 1% wage increase from Table IV. We first examine the results presented in panel A which displays the relative labour supply responses. Overall, the models predict negative and significant labour supply changes for men, but not for women. The predictions are quite robust across the three utility specifications, apart from female GPs for whom results are quite sensitive to the utility function used.

Moreover, the 5 and 10% wage increases lead to fairly linear relative changes using the translog and Box-Cox utility specifications, but less so when using the quadratic specification. A notable exception from the linear response patterns are female GPs when using the quadratic utility function, where the wage responses follow a U-shaped pattern.

[19]We use 40 hours for a standard full-time week consistent with the measure used by the National Healthcare Agreement reporting.

Panel B presents the absolute change in terms of FTE for the current population of doctors. A 5% wage increase is predicted to reduce the total number of FTE female GPs by about 0.2-1%, and for female specialists by about 0.3-0.4%, which is consistent with the modest relative wage responses by female doctors. For male GPs, a 5% (10%) increase in wages is predicted to reduce their labour supply by about 0.9-1% (1.6-1.9%). Finally, 5 and 10% wage increases would decrease the total number of FTE male specialists by 0.4-0.5% and 0.9-1.1%, respectively. That male GPs and specialists respond more strongly than female GPs and specialists is consistent with the theory of a backward bending labour supply curve and the summary statistics presented in section 5 which showed that male doctors earn higher incomes and work longer hours than female doctors. The policy simulations therefore provide evidence that wage increases in the order of 5-10% may reduce labour supply slightly in the short-run, but more so for male than female doctors.

7 Conclusion

Although the World Health Organisation (2013) predicts that most OECD (Organisation for Economic Co-operation and Development) countries will face a substantial shortage of physicians in the next years, little research exists about doctors' labour supply. We analyse the pecuniary and non-pecuniary determinants of doctors' labour supply and examine the policy implications arising from different modelling approaches for labour supply. Our study exploits the advantages of the structural discrete choice approach and compares the results to those obtained with a reduced-form approach, frequently used in the literature on physicians' labour supply.

Using a rich and unique data set on Australian physicians, "Medicine in Australia: Balancing Employment and Life" (MABEL), we make four main contributions to the literature on doctors' labour supply. First, a structural discrete choice model for labour supply is estimated for each of our four physician groups (male and female GPs and specialists). In our second contribution we make use of the rich data which allow us to perform a detailed subgroup analysis that no other study on doctors' labour supply has done before.

Third, we show that all modelling approaches used in this paper predict small negative wage elasticities for male and female GPs and specialists. Given doctors' high income levels and long working hours, it is not surprising that increasing the return to hours worked has no positive effect on their labour supply. Many doctors are working so many hours that the cost of giving up another hour of leisure is very high while the benefit from additional income is limited. As a result, wage elasticities are negative on average, although a number of individual doctors would still have a positive labour supply response. A broad range of robustness checks show that these results are robust when we account for the endogeneity of wages.

The results are not very sensitive to different definitions of non-labour and other household income, different utility specifications, different specifications of the discrete sets of working hours or including unobserved heterogeneity in preferences. This is despite the fact that for this high-income and long-working-hours population, the quadratic utility function, which is often used in general population labour supply modelling, does not fulfil coherency conditions in a large proportion of cases. The translog utility function and Box-Cox utility function have no such problem and are thus consistent with economic theory.

Finally, we use the structural model to predict relative and absolute labour supply changes in response to different wage increases. Unlike the reduced-form approach, the structural model facilitates ex ante policy simulations that explicitly take into account the non-linear taxation schedule or financial subsidies and allows for heterogeneous wage elasticities. Our policy simulations show that male doctors respond more strongly to wage increases than female doctors. A 5% increase in wages is predicted to reduce the total labour supply by male GPs by about 1%, and by about 0.5% for male specialists. Our results imply that wage increases aimed at increasing the supply of medical doctors at the intensive margin are likely to do the opposite and reduce labour supply, especially by male doctors. Our results relate to short-run impacts only since the intensive margin and not the extensive margin is investigated.

In our policy simulation we focus on the effects at the intensive margin as MABEL collects information on doctors $working$ in clinical practice, and in addition, only the intensive margin can be affected in the short run. In the long run, increased wage rates may draw in additional doctors, but given the long qualification period of doctors it is likely to take several years before any effect would be observed. Moreover, relatively few qualified doctors are likely to be absent from the medical workforce. The most notable exceptions are probably female doctors on maternity leave and recently retired doctors. These groups might respond to some extent to increased wage rates, but the net effect is ambiguous. The relevant group that is at risk of retirement (all doctors aged 55 and over) makes up 28.9% of all doctors. Higher wages may allow doctors to finance a comfortable retirement more quickly or it may incentivise doctors to stay in the workforce longer because the opportunity costs of not working as a doctor are high. This needs to be determined empirically in future research.

References

AIHW (2010). Medical labour force 2008. Technical report, Australian Institute of Health and Welfare, Canberra.

AIHW (2014). Medical workforce 2012. Report. Australian Institute of Health and Welfare.

AIHW (2015). Medical practitioner workforce. `http://www.aihw.gov.au/workforce/medical`. Australian Institute of Health and Welfare [Online; retrieved 18/03/2015].

Andreassen, L., M. L. Di Tommaso, and S. Strøm (2013). Do medical doctors respond to economic incentives? *Journal of Health Economics 32*(2), 392 – 409.

Baltagi, B. H., E. Bratberg, and T. H. Holmås (2005). A panel data study of physicians' labor supply: The case of Norway. *Health Economics 14*(10), 1035–1045.

Blundell, R. and T. MaCurdy (1999). Labor supply: A review of alternative approaches. *Handbook of Labor Economics 3*, 1559–1695.

Blundell, R., T. MaCurdy, and C. Meghir (2007). Labor supply models: unobserved heterogeneity, nonparticipation and dynamics. *Handbook of Econometrics 6*, 4667–4775.

Borjas, G. J. (1980). The relationship between wages and weekly hours of work: The role of division bias. *The Journal of Human Resources 15*(3), 409–423.

Brett, T. D., D. E. Arnold-Reed, D. A. Hince, I. K. Wood, and R. G. Moorhead (2009). Retirement intentions of general practitioners aged 45–65 years. *Medical Journal of Australia 191*(2), 75–77.

Brown, D. M. and H. E. Lapan (1972). The rising price of physicians' services: a comment. *The Review of Economics and Statistics 54*(1), 101–105.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: methods and applications*. Cambridge: Cambridge University Press.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics 18*(1), 5–46.

Cheng, T. C., G. Kalb, and A. Scott (forthcoming, 2017). Public, private or both? Analysing factors influencing the labour supply of medical specialists. *Canadian Journal of Economics*.

Cheng, T. C., A. Scott, S.-H. Jeon, G. Kalb, J. Humphreys, and C. Joyce (2012). What factors influence the earnings of general practitioners and medical specialists? Evidence from the Medicine in Australia: Balancing Employment and Life Survey. *Health Economics 21*(11), 1300–1317.

Feldstein, M. S. (1970). The rising price of physician's services. *The Review of Economics and Statistics 52*(2), 121–133.

Fuchs, V. R. and M. J. Kramer (1972). *Determinants Of Expenditures For Physicians' Services In The United States 1948-68*. Washington: United States Department of Health, Education, and Welfare.

Hanel, B., G. Kalb, and A. Scott (2014). Nurses' labour supply elasticity: The importance of accounting for extensive margins. *Journal of Health Economics 33*, 94–112.

Ikenwilo, D. and A. Scott (2007). The effects of pay and job satisfaction on the labour supply of hospital consultants. *Health Economics 16*(12), 1303–1318.

Joyce, C. M., A. Scott, S.-H. Jeon, J. Humphreys, G. Kalb, J. Witt, and A. Leahy (2010). The "Medicine in Australia: Balancing Employment and Life (MABEL)" longitudinal survey-Protocol and baseline data for a prospective cohort study of Australian doctors' workforce participation. *BMC Health Services Research 10*(1), 1–10.

Keane, M. and R. Moffitt (1998). A structural model of multiple welfare program participation and labor supply. *International Economic Review*, 553–589.

Kuehnle, D., A. Scott, T. Cheng, S. Jeon, P. Sivey, and A. Leahy (2010). Mabel user manual: Wave 1 release. *Melbourne Institute of Applied Economics and Social Research: Melbourne*.

Löffler, M., A. Peichl, and S. Siegloch (2014). Structural labor supply models and wage exogeneity. *IZA Discussion Paper Series 8281*.

Markwell, A. L. and Z. Wainer (2009). The health and wellbeing of junior doctors: insights from a national survey. *Medical Journal of Australia 191*(8), 441.

McPake, B., A. Scott, and E. Ijeoma (2014). Analyzing markets for health workers: Insights from labor and health economics. doi = http://dx.doi.org/10.1596/978-1-4648-0224-9. World Bank, Washington DC.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*(1), 69–85.

Reed, W. R. (2015). On the practice of lagging variables to avoid simultaneity. *Oxford Bulletin of Economics and Statistics 77*(6), 897–905.

Ribeiro, E. P. (2001). Asymmetric labor supply. *Empirical Economics 26*(1), 183–197.

Rizzo, J. A. and D. Blumenthal (1994). Physician labor supply: Do income effects matter? *Journal of Health Economics 13*(4), 433–453.

Sæther, E. M. (2005). Physicians' labour supply: The wage impact on hours and practice combinations. *Labour 19*(4), 673–703.

Schurer, S., D. Kuehnle, A. Scott, and T. C. Cheng (2016). A man's blessing or a woman's curse? The family earnings gap of doctors. *Industrial Relations: A Journal of Economy and Society 55*(3), 385–414.

Showalter, M. H. and N. K. Thurston (1997). Taxes and labor supply of high-income physicians. *Journal of Public Economics 66*(1), 73–97.

Shrestha, D. and C. M. Joyce (2011). Aspects of work–life balance of australian general practitioners: Determinants and possible consequences. *Australian Journal of Primary Health 17*(1), 40–47.

Sloan, F. A. (1975). Phycisian supply behavior in the short run. *Industrial & Labor Relations Review 28*, 549–569.

Stern, N. (1986). On the specification of labour supply functions. In R. Blundell and I. Walker (Eds.), *Unemployment, Search and Labour Supply*, pp. 143–189. Cambridge: Cambridge University Press.

Thornton, J. (1998). The labour supply behaviour of self-employed solo practice physicians. *Applied Economics 30*(1), 85–94.

Van Soest, A. (1995). Structural models of family labor supply: a discrete choice approach. *Journal of Human Resources*, 63–88.

Van Soest, A., M. Das, and X. Gong (2002). A structural labour supply model with flexible preferences. *Journal of Econometrics 107*(1), 345–374.

Van Soest, A., I. Woittiez, and A. Kapteyn (1990). Labor supply, income taxes, and hours restrictions in the netherlands. *Journal of Human Resources*, 517–558.

Varian, H. R. (1992). *Microeconomic analysis*, Volume 2. Norton New York.

Wang, C. and A. Sweetman (2013). Gender, family status and physician labour supply. *Social Science & Medicine 94*, 17–25.

Wooldridge, J. M. (2010). Correlated random effects models with unbalanced panels. Working Paper. Michigan State University, Department of Economics.

World Health Organisation (2013). A universal truth: No health without a workforce. Report. World Health Organisation.

Figure 1: Percentage of female doctors in the medical workforce and average weekly hours worked by age group, gender and doctor type

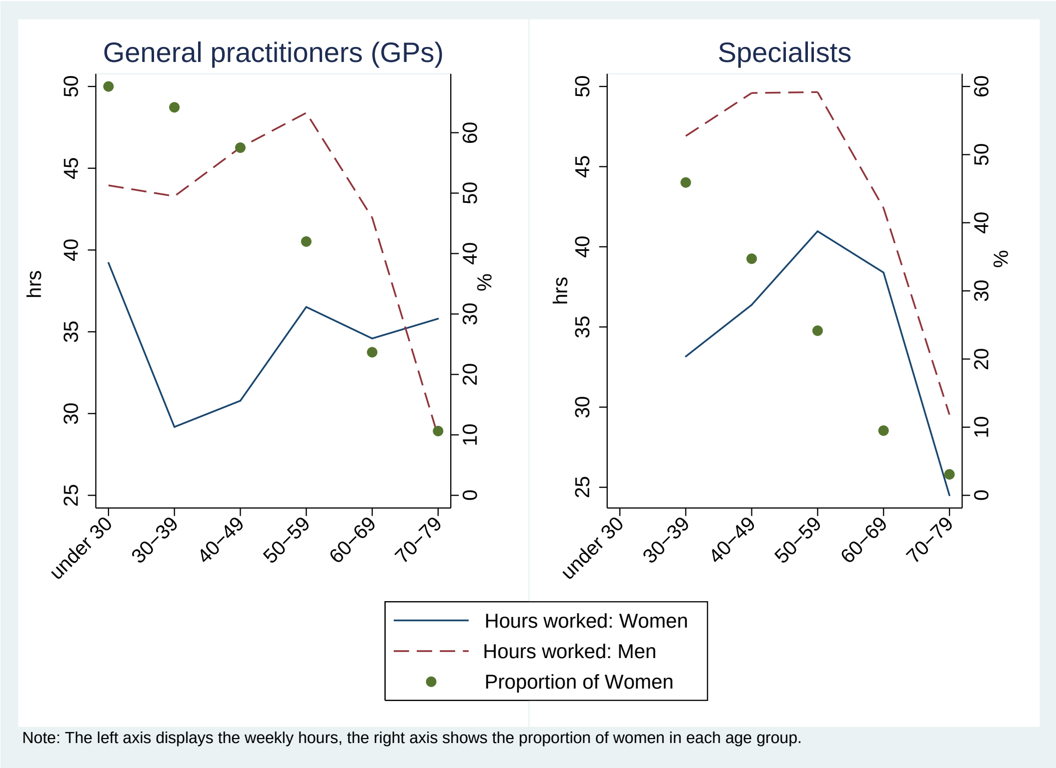

Figure 2: Kernel density distribution of hours worked

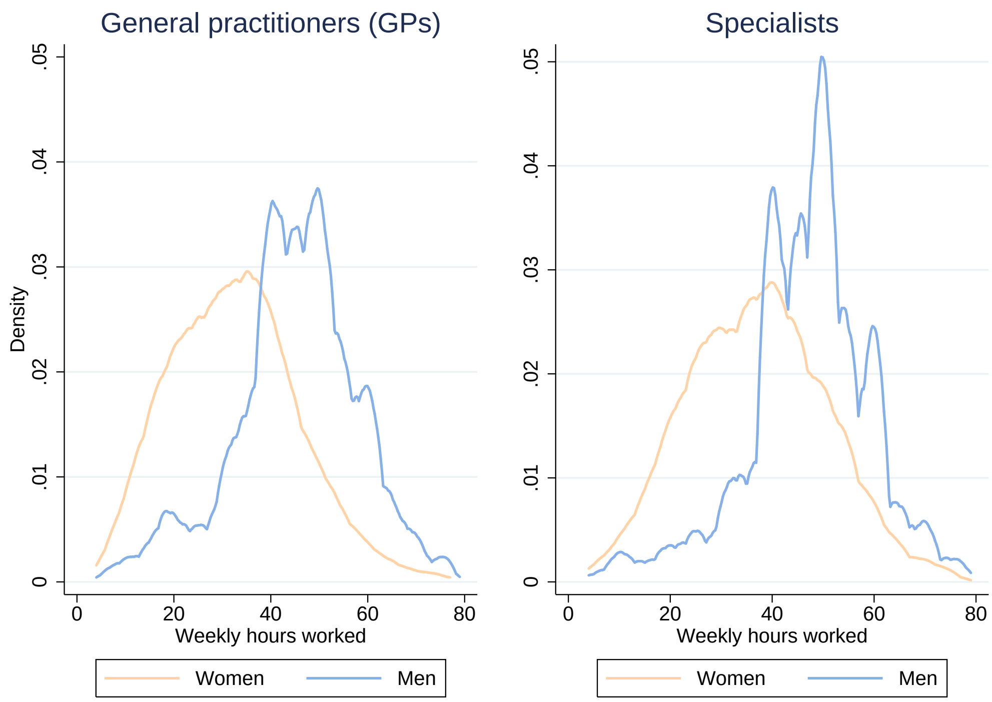

Figure 3: Distribution of observed hourly wages, by age group, gender and doctor type

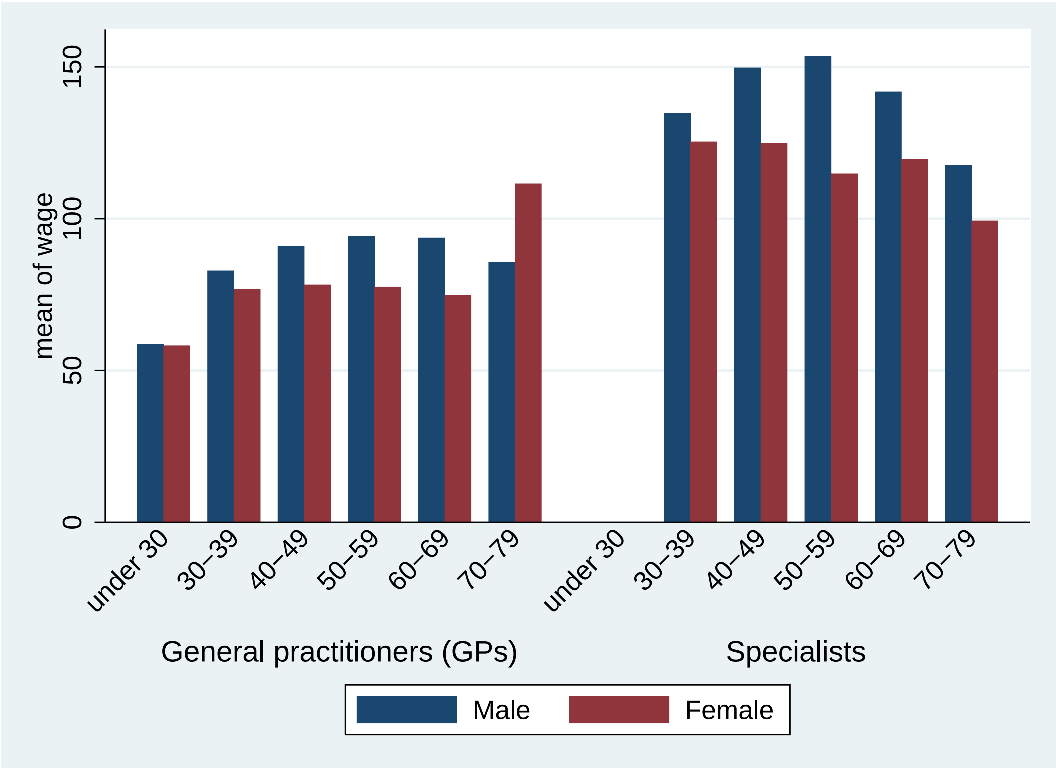

Figure 4: Distribution of wage elasticities across individual doctors (imputed wages, 10 mid-points, translog utility function)

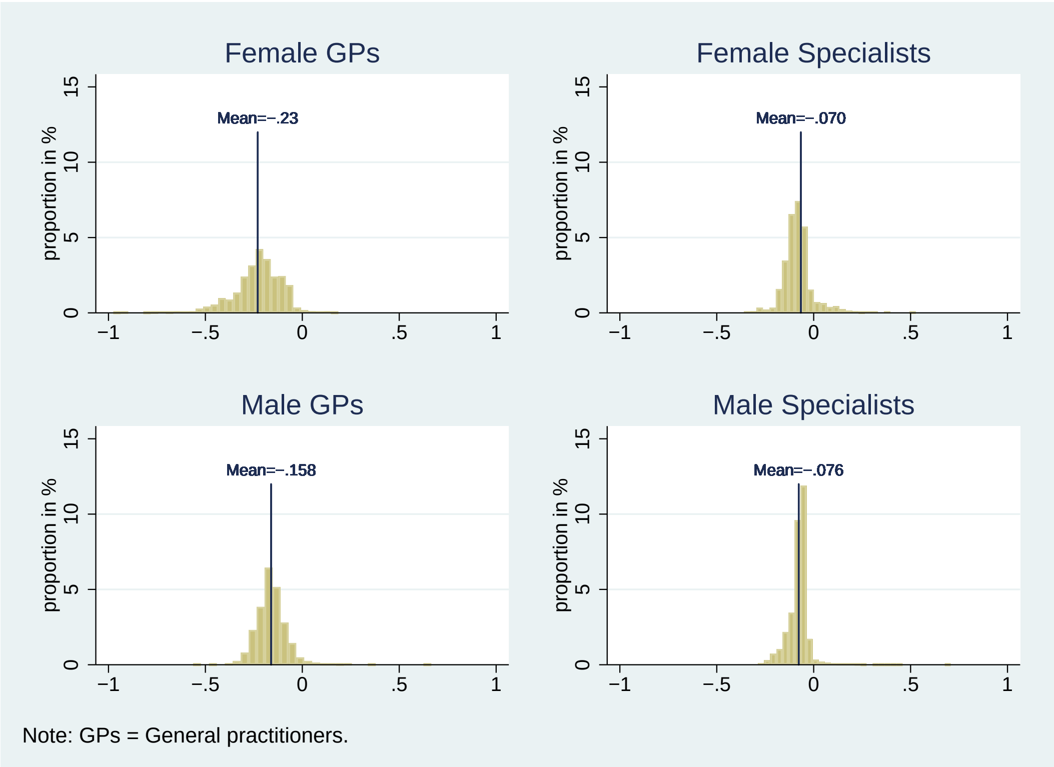

Figure 5: Estimated wage elasticities for subgroups, by doctor type and gender (imputed wages, 10 mid-points, translog utility function)

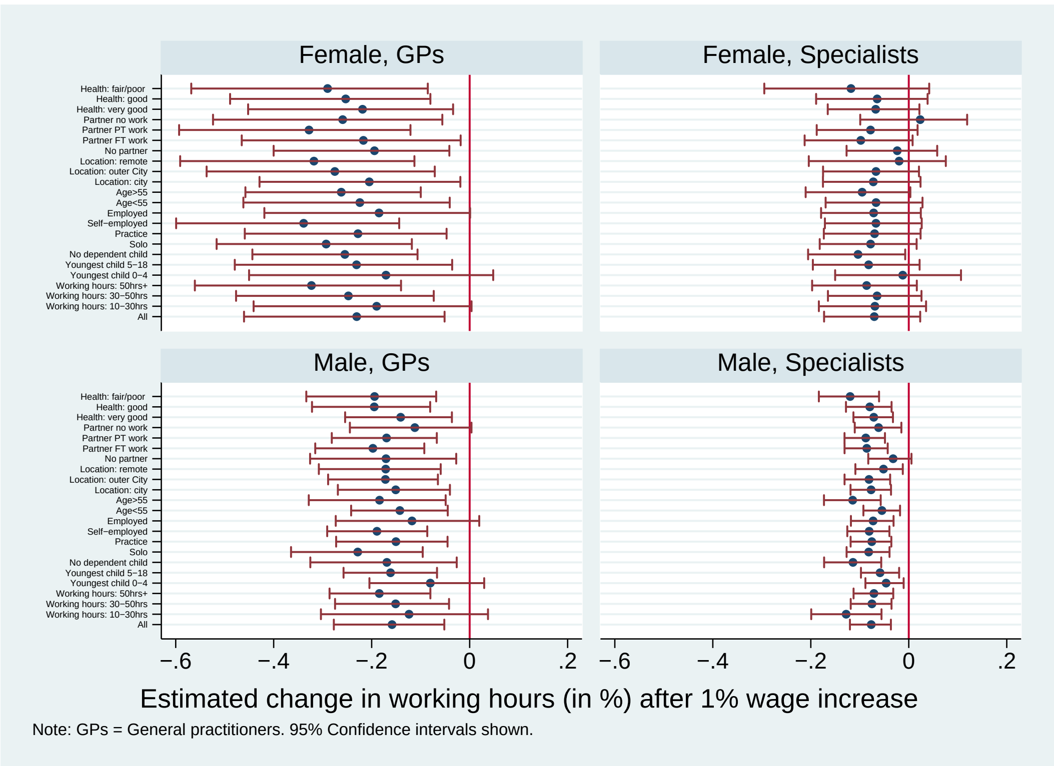

Table I: Summary statistics by gender and doctor type

	Female				Male			
	GPs		Specialists		GPs		Specialists	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Weekly net income in $	1755.4	(830.4)	2843.1	(1625.6)	2668.4	(1179.1)	4178.5	(2261)
Weekly hours	32.5	(13.0)	36.8	(13.4)	45.1	(12.7)	47.1	(11.8)
Hourly gross wage in $	76.6	(32.4)	122.5	(69.4)	91.2	(41.4)	146.8	(81.4)
Age/10	4.6	(0.9)	4.6	(0.8)	5.2	(1.0)	5.1	(1.0)
No children/youngest child over 15	0.283	(0.5)	0.299	(0.5)	0.346	(0.5)	0.320	(0.5)
Number of dependent children (under 25)	1.600	(1.3)	1.493	(1.2)	1.509	(1.4)	1.629	(1.4)
Youngest child 0-4	0.174	(0.4)	0.252	(0.4)	0.117	(0.3)	0.167	(0.4)
Youngest child 5-9	0.154	(0.4)	0.160	(0.4)	0.113	(0.3)	0.146	(0.4)
Youngest child 10-15	0.206	(0.4)	0.161	(0.4)	0.169	(0.4)	0.176	(0.4)
No partner	0.133	(0.3)	0.178	(0.4)	0.072	(0.3)	0.051	(0.2)
Partner	0.867	(0.3)	0.822	(0.4)	0.928	(0.3)	0.949	(0.2)
Partner works	0.769	(0.4)	0.730	(0.4)	0.624	(0.5)	0.647	(0.5)
Partner works full-time	0.657	(0.5)	0.576	(0.5)	0.226	(0.4)	0.205	(0.4)
Partner works part-time	0.112	(0.3)	0.153	(0.4)	0.398	(0.5)	0.442	(0.5)
Partner does not work	0.097	(0.3)	0.092	(0.3)	0.304	(0.5)	0.302	(0.5)
Self-employed	0.296	(0.5)	0.273	(0.4)	0.570	(0.5)	0.468	(0.5)
Employed	0.704	(0.5)	0.727	(0.4)	0.430	(0.5)	0.532	(0.5)
Very good health	0.735	(0.4)	0.744	(0.4)	0.671	(0.4)	0.724	(0.4)
Good health	0.191	(0.3)	0.187	(0.3)	0.214	(0.4)	0.203	(0.4)
Fair/poor health	0.074	(0.2)	0.069	(0.2)	0.115	(0.3)	0.073	(0.2)
City	0.705	(0.5)	0.882	(0.3)	0.636	(0.5)	0.824	(0.4)
Outer city	0.180	(0.4)	0.090	(0.3)	0.226	(0.4)	0.140	(0.3)
Remote	0.115	(0.3)	0.029	(0.2)	0.138	(0.3)	0.036	(0.2)
NSW	0.259	(0.4)	0.282	(0.4)	0.262	(0.4)	0.299	(0.4)
ACT	0.026	(0.2)	0.013	(0.1)	0.013	(0.1)	0.017	(0.1)
NT	0.007	(0.1)	0.007	(0.1)	0.011	(0.1)	0.007	(0.1)
QLD	0.205	(0.4)	0.169	(0.4)	0.192	(0.4)	0.169	(0.4)
SA	0.071	(0.3)	0.113	(0.3)	0.100	(0.3)	0.084	(0.3)
TAS	0.043	(0.2)	0.029	(0.2)	0.035	(0.2)	0.027	(0.2)
VIC	0.283	(0.5)	0.319	(0.5)	0.279	(0.4)	0.317	(0.5)
WA	0.106	(0.3)	0.069	(0.3)	0.107	(0.3)	0.081	(0.3)
Number of observations	1067		769		1128		1908	

Table II: Comparison of goodness of fit for different utility functions.

Women	GPs			Specialist		
	Quadratic	Translog	Box-Cox[a]	Quadratic	Translog	Box-Cox
% quasi-concave	17	91	100	57	98	97
% where U increases with y	28	97	100	80	100	98
% where indifference curve convex	36	94	100	72	98	98
Average predicted probability at observed LS point	21.6%	21.5%	20.3%	17.6%	18.0%	17.9%
% correctly predicted	33.3%	30.6%	29.8%	27.8%	29.5%	28.6%
Average predicted hours	33.3	33.3	33.3	37.2	37.2	37.2
Average observed hours	32.5	32.5	32.5	36.8	36.8	36.8
parameters	33	33	18	33	33	33
loglikelihood value	-1838	-1856	-1894	-1457	-1444	-1447
Akaike Information Criterion (AIC)	3742	3777	3824	2979	2954	2960
Bayesian Information Criterion (BIC)	3982	4017	3955	3209	3183	3189
Number of observations		1067			769	

Men	GPs			Specialists		
	Quadratic	Translog	Box-Cox	Quadratic	Translog	Box-Cox
% quasi-concave	15	95	89	23	99	99
% where U increases with y	31	97	93	23	100	99
% where indifference curve convex	37	98	94	23	99	100
Average predicted probability at observed LS point	16.3%	16.5%	16.5%	15.6%	15.7%	15.7%
% correctly predicted	24.6%	25.5%	25.7%	21.8%	20.4%	20.5%
Average predicted hours	45.4	45.4	45.4	47.3	47.2	47.2
Average observed hours	45.1	45.1	45.1	47.1	47.1	47.1
parameters	33	33	33	33	33	33
loglikelihood value	-2199	-2195	-2193	-3772	-3768	-3762
AIC	4463	4456	4452	7609	7601	7590
BIC	4705	4698	4694	7869	7860	7849
Number of observations		1128			1908	

Notes: Shaded cells indicate the best performing model on the criterion described in the first column of that row.
[a] Model estimated without interaction terms between income and hours worked for this group.

Table III: Marginal effects measured as an absolute change in hours worked for the labour supply model with 10 discrete points, translog utility function, imputed wages

Panel A: Women	GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-11.52	[-13.70, -8.73]	-10.82	[-13.23, -8.15]
5-9	-8.33	[-10.49, -5.54]	-4.87	[-8.00, -0.57]
10-15	-3.96	[-6.11, -1.80]	-0.20	[-3.26, 3.69]
Number of children	-0.66	[-1.32, 0.02]	-0.78	[-1.76, 0.12]
Age	-0.16	[-0.25, -0.07]	0.00	[-0.13, 0.15]
Health status (ref. group: very good)				
good health	3.27	[1.59, 5.08]	0.19	[-1.82, 2.46]
poor/fair health	1.73	[-0.73, 4.72]	1.16	[-2.11, 5.05]
Partnership status (ref. group: single)				
Full-time work	-3.80	[-5.91, -1.53]	-4.39	[-6.73, -1.93]
Part-time work	-1.33	[-3.93, 1.68]	-2.44	[-5.31, 0.78]
Not employed	4.01	[1.00, 7.18]	3.97	[0.91, 6.96]
Self-employed	9.19	[7.22, 12.01]	4.88	[2.82, 7.15]
Location (ref. group: urban)				
Inner regional	1.47	[-0.24, 3.34]	1.20	[-1.64, 4.10]
Remote	6.73	[4.64, 9.20]	1.61	[-3.75, 7.28]
Panel B: Men	**GPs**		**Specialists**	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-4.45	[-7.25, -1.76]	-1.53	[-3.94, 0.64]
5-9	-2.19	[-4.71, 0.06]	-0.78	[-2.72, 1.14]
10-15	-1.95	[-4.32, 0.23]	-0.31	[-2.24, 1.32]
Number of children	1.39	[0.78, 2.00]	0.93	[0.43, 1.42]
Age	-0.17	[-0.25, -0.09]	-0.23	[-0.30, -0.17]
Health status (ref. group: very good)				
good health	2.19	[0.64, 3.71]	-0.13	[-1.32, 1.02]
poor/fair health	0.48	[-1.63, 2.55]	0.50	[-1.45, 2.39]
Partnership status (ref. group: single)				
Full-time work	0.29	[-2.68, 3.17]	-0.96	[-2.94, 1.02]
Part-time work	-1.31	[-4.43, 1.57]	-1.41	[-3.41, 0.61]
Not employed	-0.16	[-3.00, 2.84]	-0.32	[-2.41, 1.63]
Self-employed	7.60	[6.40, 9.15]	3.56	[2.48, 4.72]
Location (ref. group: urban)				
Inner regional	1.53	[-0.14, 3.05]	-0.30	[-1.75, 1.07]
Remote	4.20	[2.25, 6.17]	-0.05	[-2.34, 2.07]

Table IV: Comparison of simulated wage elasticities

	Women				Men			
	GPs		Specialists		GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs	Point est.	95% CIs	Point est.	95% CIs
Panel A: structural model (translog specification)								
Using observed wage								
5 mid-points	-0.086	[-0.14, -0.04]	-0.086	[-0.14, -0.04]	-0.076	[-0.11, -0.05]	-0.097	[-0.12, -0.08]
10 mid-points	-0.102	[-0.15, -0.06]	-0.076	[-0.12, -0.03]	-0.081	[-0.11, -0.05]	-0.107	[-0.13, -0.09]
13 mid-points	-0.111	[-0.17, -0.06]	-0.080	[-0.13, -0.04]	-0.079	[-0.11, -0.05]	-0.107	[-0.13, -0.09]
Using imputed wages								
5 mid-points	-0.081	[-0.31, 0.11]	-0.101	[-0.21, 0.00]	-0.153	[-0.28, -0.04]	-0.066	[-0.11, -0.03]
10 mid-points	-0.230	[-0.46, -0.05]	-0.070	[-0.17, 0.02]	-0.158	[-0.28, -0.05]	-0.076	[-0.12, -0.04]
13 mid-points	-0.226	[-0.43, -0.03]	-0.065	[-0.17, 0.03]	-0.145	[-0.26, -0.04]	-0.061	[-0.11, -0.02]
Panel B: structural model (imputed wage, 10 midpoints)								
Quadratic	-0.123	[-0.300, 0.053]	-0.073	[-0.172, 0.018]	-0.176	[-0.294, -0.060]	-0.093	[-0.137, -0.050]
Reduced Box-Cox[a]	-0.018	[-0.146, 0.141]	-0.030	[-0.110, 0.069]	-0.020	[-0.087, 0.059]	-0.023	[-0.054, 0.014]
Full Box-Cox[b]	d.n.c.	[d.n.c.]	-0.088	[-0.236, 0.005]	-0.185	[-0.304, -0.012]	-0.077	[-0.120, -0.004]
Panel C: reduced-form model								
Observed wage	-0.105	[-0.171, -0.040]	-0.103	[-0.161, -0.045]	-0.113	[-0.153, -0.074]	-0.134	[-0.164, -0.103]
Imputed wage[c]	-0.056	[-0.311, 0.207]	-0.070	[-0.185, 0.046]	-0.202	[-0.339, -0.065]	-0.088	[-0.146, -0.028]
Observed wage (IV1)[d]	-0.064	[-0.329, 0.199]	-0.080	[-0.194, 0.035]	-0.200	[-0.341, -0.054]	-0.108	[-0.165, -0.052]
Number of observations	1067		769		1128		1908	

Notes: [a]*: The reduced Box-Cox function does not include an interaction term between income and hours worked.*

[b]*: The full Box-Cox function includes an interaction term between income and hours worked as in equation* (3).

[c]*: We obtain the imputed wages from the wage regressions presented in Table A.I.*

[d]*: In the IV regressions we control for the same variables as in Table A.VII, and instrument for the wage using the equation presented in TableA.I.*

d.n.c. = model did not converge.

Table V: Policy simulation: changes in working hours due to different wage increases (10 mid-points, imputed wages).

	Women				Men			
	GPs		Specialists		GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs	Point est.	95% CIs	Point est.	95% CIs
Panel A: Predicted relative changes (%) in hours worked in response to simulated wage increases								
1% wage increase								
Quadratic	-0.123	[-0.300,0.053]	-0.073	[-0.172,0.018]	-0.176	[-0.294,-0.060]	-0.093	[-0.137,-0.050]
Translog	-0.231	[-0.461,-0.051]	-0.070	[-0.173,0.023]	-0.158	[-0.277,-0.052]	-0.077	[-0.120,-0.037]
Box Cox	-0.018[a]	[-0.146,0.141]	-0.088	[-0.236,0.005]	-0.185	[-0.304,-0.012]	-0.077	[-0.119, -0.004]
5% wage increase								
Quadratic	-0.370	[-1.280,0.572]	-0.276	[-0.770,0.180]	-0.787	[-1.404,-0.193]	-0.470	[-0.694,-0.251]
Translog	-1.076	[-2.150,-0.221]	-0.357	[-0.857,0.101]	-0.748	[-1.334,-0.230]	-0.383	[-0.597,-0.187]
Box Cox	-0.082[a]	[-0.708,0.703]	-0.442	[-1.150,0.012]	-0.893	[-1.469,-0.049]	-0.380	[-0.589,-0.022]
10% wage increase								
Quadratic	-0.181	[-2.163,1.870]	-0.325	[-1.357,0.658]	-1.340	[-2.747,-0.049]	-0.958	[-1.427,-0.505]
Translog	-1.988	[-3.994,-0.351]	-0.692	[-1.664,0.197]	-1.407	[-2.536,-0.415]	-0.757	[-1.182,-0.371]
Box Cox	-0.155[a]	[-1.373,1.385]	-0.866	[-2.219,0.024]	-1.725	[-2.822,-0.088]	-0.745	[-1.151,-0.046]
Panel B: Predicted absolute changes (in FTE workers) in response to simulated wage increases[b]								
1% wage increase								
Quadratic	-8.07	[-22.36,5.53]	-4.66	[-10.38,0.60]	-31.07	[-49.93,-12.20]	-16.85	[-25.07,-9.04]
Translog	-19.14	[-37.81,-5.53]	-4.21	[-9.93,0.90]	-27.74	[-46.23,-9.99]	-14.38	[-22.60,-6.99]
Box Cox	-1.61[a]	[-11.07,10.61]	-5.01	[-11.68,0.11]	-30.73	[-48.52,-2.37]	-14.63	[-22.77,-1.03]
5% wage increase								
Quadratic	-22.13	[-93.14,49.57]	-18.51	[-46.50,7.07]	-139.79	[-239.28,-42.16]	-86.30	[-126.99,-44.80]
Translog	-89.68	[-173.83,-23.98]	-21.22	[-48.60,3.76]	-130.55	[-221.90,-45.49]	-72.74	[-113.02,-36.17]
Box Cox	-6.92[a]	[-54.18,52.57]	-25.14	[-57.01,0.09]	-148.48	[-234.76,-10.13]	-72.66	[-112.40,-5.71]
10% wage increase								
Quadratic	-2.31	[-157.93,160.92]	-24.38	[-82.01,29.64]	-241.50	[-453.04,-29.96]	-175.90	[-260.15,-88.36]
Translog	-165.53	[-321.62,-43.11]	-40.93	[-93.90,7.98]	-246.30	[-428.26,-83.58]	-143.84	[-221.52,-72.33]
Box Cox	-13.14[a]	[-104.90,103.75]	-49.16	[-110.49,-0.21]	-287.10	[-452.59,-17.79]	-142.40	[-219.54,-11.63]

Note: [a] *For female GPs, a Box Cox model without an income and labour supply interaction term is used for the simulation.*
[b] *Absolute changes are obtained by multiplying relative changes by the number of doctors according to the AIHW (2010). In 2008, there were 9,222 female GPs, 14,793 male GPs, 6,019 female specialists and 16,439 male specialists.*

Online Appendix to Kalb et al. "What factors affect physicians labour supply: Comparing structural discrete choice and reduced-form approaches"

Table A.I: OLS of ln(wage)

	Women		Men	
	GPs	Specialists	GPs	Specialists
Australian medical school	-0.096***	-0.111**	0.010	0.005
	(0.036)	(0.047)	(0.037)	(0.028)
Number of postgraduate qualifications	-0.026	-0.030	0.071	-0.017
	(0.042)	(0.098)	(0.047)	(0.066)
Temporary visa holder	0.037	-0.057	-0.084	-0.122
	(0.107)	(0.213)	(0.091)	(0.131)
Actual work experience				
15-19 years	-0.025	0.073	0.173***	0.091**
	(0.039)	(0.047)	(0.058)	(0.040)
20-24 years	0.005	0.046	0.062	0.125***
	(0.036)	(0.049)	(0.050)	(0.039)
25-29 years	-0.032	-0.046	0.143***	0.093**
	(0.039)	(0.052)	(0.048)	(0.039)
30-34 years	-0.055	-0.006	0.087*	0.083**
	(0.049)	(0.063)	(0.049)	(0.041)
35-39 years	-0.030	0.043	0.043	0.052
	(0.071)	(0.106)	(0.057)	(0.043)
40-45 years	0.164	0.004	0.103	0.051
	(0.111)	(0.158)	(0.069)	(0.049)
45 or more years	-0.108	-0.147	0.097	-0.178***
	(0.286)	(0.263)	(0.078)	(0.059)
State dummies				
VIC	-0.041	0.028	0.068*	-0.032
	(0.034)	(0.043)	(0.038)	(0.026)
QLD	0.044	0.177***	0.090**	0.146***
	(0.038)	(0.054)	(0.043)	(0.032)
SA	-0.012	0.065	0.038	-0.012
	(0.054)	(0.060)	(0.053)	(0.041)
WA	0.022	0.063	0.146***	0.057
	(0.047)	(0.071)	(0.052)	(0.041)
NT	0.017	0.382	0.007	-0.081
	(0.160)	(0.239)	(0.143)	(0.140)
TAS	-0.033	0.004	-0.020	-0.155**
	(0.069)	(0.118)	(0.081)	(0.070)
ACT	0.042	0.071	0.043	-0.031
	(0.080)	(0.147)	(0.123)	(0.081)
Inner regional area	-0.016	-0.051	0.070*	0.067*
	(0.038)	(0.072)	(0.040)	(0.035)
Remote area	0.051	-0.223*	0.104**	0.044
	(0.045)	(0.118)	(0.049)	(0.064)
Self-employed	0.163***	0.139***	0.067**	0.190***
	(0.029)	(0.041)	(0.031)	(0.024)
Practice size				
2-3 doctors	-0.024		0.147***	
	(0.069)		(0.053)	
4-5 doctors	0.020		0.180***	
	(0.068)		(0.053)	

...table A.I continued

	Women		Men	
	GPs	Specialists	GPs	Specialists
6-9 doctors	-0.006		0.237***	
	(0.068)		(0.050)	
10 or more doctors	0.069		0.340***	
	(0.072)		(0.058)	
PG Certificate or Diploma	0.040	0.004	-0.098	0.002
	(0.058)	(0.125)	(0.065)	(0.083)
Masters or PhD	-0.029	-0.001	-0.128	-0.003
	(0.069)	(0.116)	(0.084)	(0.083)
Fellowship of Colleges	0.076***	0.047	0.023	0.051
	(0.028)	(0.101)	(0.030)	(0.056)
Other qualifications	0.111	-0.025	-0.105	0.039
	(0.086)	(0.135)	(0.088)	(0.091)
% of time in clinical work	0.003***	0.003***	0.001	0.004***
	(0.001)	(0.001)	(0.001)	(0.000)
Local median house price	0.009**	-0.002	0.009*	0.009***
(in $\$100,000$)	(0.004)	(0.004)	(0.005)	(0.003)
Main speciality				
Cardiology		0.342		-0.095
		(0.470)		(0.107)
Gastroenterology		0.577		-0.079
		(0.462)		(0.097)
General medicine		0.464		-0.286***
		(0.458)		(0.096)
Intensive care internal medicine		0.789		
		(0.502)		
Paediatric medicine		0.259		-0.372***
		(0.451)		(0.087)
Thoracic medicine		0.013		-0.233**
		(0.467)		(0.101)
Other internal medicine		0.410		-0.244***
		(0.450)		(0.079)
Pathology		0.687		-0.004
		(0.453)		(0.090)
General surgery		0.402		-0.000
		(0.464)		(0.087)
Orthopaedic surgery		0.952*		0.236***
		(0.517)		(0.091)
Other surgery		0.427		0.156*
		(0.456)		(0.087)
Anaesthesia		0.740*		0.084
		(0.449)		(0.078)
Diagnostic radiology		0.761*		0.273***
		(0.455)		(0.087)
Obstetrics and gynaecology		0.757*		0.079
		(0.452)		(0.087)
Psychiatry		0.411		-0.239***
		(0.450)		(0.080)
Number of observations	1067	769	1128	1908

*Note: Significance is indicated with * for 10% level, ** for 5% level and *** for 1% level.*

Additional explanation and discussion of Table A.II

First, Panel A presents the results for the structural model when we instrument current wages using lagged wages. This estimation is based on the third wave of the MABEL data and uses lagged observed wages from waves 1 and 2 as exclusion restrictions for observed wages in wave 3. The results are similar to the results from our main specifications.

Second, Panel B presents the results using imputed wages which are obtained from the wage regressions applying a correlated random effects approach using four waves of data (see equation 8). The results are almost identical to our main specification.

Third, similar to Van Soest (1995), we also estimate the model taking into account errors in wage rate predictions by drawing 100 wages for each individual, using the standard deviation of the wage regressions. Panel C in Table A.II shows that the estimated wage elasticities for three of the doctor groups are robust to allowing for wage rate prediction errors. However, the model does again not converge for female GPs.

Fourth, Panel D presents the results for alternative wage definitions. The results are qualitatively the same across specifications with the largest differences observed for female GPs.

Fifth, Individuals who are most likely to face demand side factors that lead to sub-optimal working hours are those for whom observed hours are not equal to preferred hours. Since preferred hours are seen as a reflection of actual unrestricted labour supply, this may potentially lead to bias in the estimation of the model's parameters due to measurement error. Therefore, we estimate an alternative version of the model, excluding all observations who are not working at their preferred hours, following Ribeiro (2001) who uses information from the sample (whether workers were looking for another job) to exclude individuals from the analysis. This provides an indication of the bias of the estimated elasticities due to sub-optimal labour supply reported in the data. Unfortunately, the question in MABEL is not ideal since it is not asked conditional on income changing with a change in hours worked, but the results provide some indication to the sensitivity of our elasticity to leaving out doctors who state they would like to change hours worked. After dropping these individuals from the analysis, the estimation results remain of the same order of magnitude, although they become mostly insignificant due to the much smaller sample size. These results are presented in panel E.

Sixth, we also estimate separate models for those who are employed versus those who are self-employed. Interestingly, the results (see Panel F) are not qualitatively different (except for female GPs) with all elasticities varying between -0.16 and -0.03. The proportion of female doctors who are self employed is much lower than for male doctors (see Table I) leading to small subsample sizes; so that for female GPs (and to a lesser extent female specialists), this self-employed group has much wider confidence intervals around the point estimates of wage elasticity than the male doctor types.

Table A.II: Comparison of simulated wage elasticities, translog benchmark specification (10 mid-points)

	Women				Men			
	GPs		Specialists		GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs	Point est.	95% CIs	Point est.	95% CIs
Panel A: using lagged wages as instruments for current wages								
Imputed wage based on lagged wages	-0.045	[-0.175, 0.054]	-0.097	[-0.196, -0.014]	-0.048	[-0.125, 0.025]	-0.079	[-0.116, -0.048]
Number of observations	*524*		*394*		*496*		*987*	
Panel B: using four waves of wage data to impute wages								
Imputed wage based on four waves [a]	-0.208	[-0.424, -0.030]	-0.068	[-0.160, 0.018]	-0.165	[-0.279, -0.068]	-0.084	[-0.126, -0.044]
Number of person-year observations	*2688*		*2074*		*2697*		*5170*	
Panel C: accounting for wage prediction error								
using 100 wage draws	d.n.c.[b]	[d.n.c.]	-0.083	[-0.178, 0.017]	-0.173	[-0.246, -0.083]	-0.085	[-0.127, -0.042]
Number of observations	*1067*		*769*		*1128*		*1908*	
Panel D: Alternative wage specifications								
Alternative wage definition 1[c]	-0.079	[-0.139, -0.027]	-0.079	[-0.140, -0.026]	-0.079	[-0.122, -0.045]	-0.113	[-0.144, -0.09]
Number of observations	*803*		*559*		*782*		*1215*	
Alternative wage definition 2[d]	-0.105	[-0.158, -0.060]	-0.092	[-0.140, -0.049]	-0.086	[-0.118, -0.059]	-0.109	[-0.130, -0.091]
Number of observations	*1107*		*790*		*1149*		*1929*	
Alternative wage definition 3[e]	-0.045	[-0.090, -0.004]	-0.084	[-0.127, -0.045]	-0.081	[-0.108, -0.056]	-0.096	[-0.116, -0.080]
Number of observations	*1145*		*821*		*1173*		*1963*	
Panel E: excluding those who are not at their preferred hours of work								
excl. mismatches	-0.058	[-0.337, 0.226]	0.109	[-0.064, 0.278]	-0.228	[-0.469, -0.034]	-0.017	[-0.093, 0.057]
Number of observations	*584*		*350*		*444*		*739*	
Panel F: separate estimations for employed and self-employed doctors								
employed	-0.272	[-0.530, -0.014]	-0.030	[-0.165, 0.092]	-0.131	[-0.360, 0.066]	-0.104	[-0.166, -0.047]
Number of observations	*751*		*559*		*485*		*1,016*	
self-employed	0.007	[-0.418, 0.424]	-0.136	[-0.337, 0.046]	-0.160	[-0.301, -0.009]	-0.047	[-0.112, 0.020]
Number of observations	*316*		*210*		*643*		*892*	

Notes: [a]*: These imputed wages are obtained from wage regressions (as described in equation 8) which apply a correlated random effects approach using four waves of data.*
[b] *d.n.c. = model did not converge.*
[c] *This definition uses the proportion of income doctors earned through medical practice and through other sources to assign other household income to the doctor and the doctor's partner. The sample is reduced by about 25% when using this definition.*
[d] *This definition also includes doctors reporting net income only by applying tax and transfer rules in reverse to obtain gross income. Otherwise the same approach to assign income to the doctor and the doctor's partner is the same as under the main approach. An additional 282 doctors can be included.*
[e] *In this definition, the previous two approaches are combined.*

Table A.III: Coefficients from multinomial logit model with 10 points, translog utility function, imputed wages

	Women				Men			
	GPs		Specialists		GPs		Specialists	
	coef	S.E.	coef	S.E.	coef	S.E.	coef	S.E.
Weekly net income	-54.725**	(26.177)	-14.561	(17.954)	-46.763**	(22.576)	18.077	(12.251)
Weekly net income^2	2.210*	(1.174)	0.349	(0.717)	1.692	(1.172)	-0.789	(0.572)
Weekly net income interacted with								
Weekly hours	8.325***	(2.266)	2.349*	(1.259)	5.283***	(1.485)	0.814	(0.684)
Age of youngest child (reference group: no child)								
0-4	-5.320***	(0.971)	-2.768**	(1.258)	0.321	(1.433)	-0.393	(1.208)
5-9	-3.936***	(0.902)	-3.487***	(1.141)	1.154	(1.436)	0.605	(1.159)
10-15	-1.190	(0.834)	-2.249**	(1.067)	-0.308	(1.108)	0.774	(1.044)
Age	-0.198	(2.926)	4.476	(4.107)	3.552	(2.294)	2.380	(2.498)
Age squared	-0.037	(0.311)	-0.514	(0.424)	-0.472**	(0.210)	-0.371*	(0.223)
Number of children	-0.257	(0.256)	0.899**	(0.375)	0.334	(0.341)	0.693**	(0.296)
Partner's employment (reference group: single)								
not employed	0.028	(1.416)	1.141	(1.522)	1.808	(1.229)	-3.558**	(1.513)
works part-time	-2.482**	(1.189)	-1.870	(1.170)	0.813	(1.179)	-3.584**	(1.501)
works full-time	-2.339**	(1.042)	-3.102***	(0.993)	0.778	(1.215)	-3.436**	(1.559)
Self-employed	0.467	(0.735)	0.680	(0.733)	1.244	(0.768)	0.673	(0.600)
Location (reference group: urban)								
Outer city	0.147	(0.681)	1.293	(1.068)	-0.311	(0.759)	-0.405	(0.722)
Remote	1.486*	(0.834)	1.013	(1.941)	-0.299	(0.945)	2.233	(1.580)
Health (reference group: very good)								
good health	1.209*	(0.682)	-0.428	(0.739)	-0.666	(0.729)	-0.281	(0.620)
fair/poor health	-0.054	(0.978)	-2.396**	(1.013)	-1.451*	(0.865)	-1.842**	(0.824)
Weekly hours	-80.061***	(22.099)	-26.945*	(15.218)	-44.786***	(14.334)	2.171	(7.814)
Weekly hours^2	4.116***	(0.753)	2.583***	(0.593)	1.537***	(0.567)	1.234***	(0.345)
Weekly working hours interacted with								
Age of youngest child (reference group: no child)								
0-4	-1.675	(1.029)	0.759	(1.364)	1.591	(0.986)	0.176	(0.669)
5-9	-1.298	(0.891)	-1.818*	(1.070)	1.425*	(0.846)	0.587	(0.619)
10-15	0.165	(0.818)	-1.773**	(0.892)	0.353	(0.616)	0.548	(0.545)
Age	-2.019	(2.899)	-0.980	(3.766)	-0.711	(1.868)	-3.832**	(1.751)
Age squared	0.216	(0.309)	0.079	(0.388)	0.021	(0.175)	0.327**	(0.163)
Number of children	-0.025	(0.239)	1.089***	(0.363)	-0.190	(0.188)	0.145	(0.158)
Partner's employment (reference group: single)								
not employed	-1.300	(1.116)	-0.409	(1.071)	1.189	(0.821)	-1.957**	(0.893)
works part-time	-1.912*	(1.036)	-0.886	(0.966)	0.885	(0.781)	-1.674*	(0.883)
works full-time	-0.962	(0.912)	-1.485*	(0.777)	0.408	(0.793)	-1.708*	(0.910)
Self-employed	-2.328***	(0.708)	-0.736	(0.623)	-1.407***	(0.519)	-0.586	(0.364)
Location (reference group: urban)								
Outer city	-0.374	(0.655)	0.811	(0.942)	-0.642	(0.462)	-0.165	(0.426)
Remote	-0.783	(0.717)	0.437	(1.521)	-1.341**	(0.566)	1.366	(0.961)
Health (reference group: very good)								
good health	0.055	(0.620)	-0.447	(0.647)	-1.039**	(0.449)	-0.134	(0.374)
fair/poor health	-0.628	(0.874)	-2.340***	(0.767)	-1.084**	(0.551)	-1.230**	(0.489)
Number of observations	1067		769		1128		1908	

Note: for ease of reporting, weekly net income has been divided by 1000, and weekly hours and age have been divided by 10.
*Significance is indicated with * for 10% level, ** for 5% level and *** for 1% level.*

Table A.IV: Marginal effects on hours worked for labour supply model with 10 discrete points, quadratic utility function, imputed wages

Panel A: Women	GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-12.15	[-14.10, -10.10]	-11.24	[-13.93, -8.37]
5-9	-9.20	[-11.14, -7.11]	-6.49	[-9.61, -3.43]
10-15	-4.59	[-6.73, -2.73]	-1.60	[-4.48, 1.16]
Number of children	-1.37	[-2.48, -0.03]	-0.96	[-2.77, 1.02]
Age	-0.14	[-0.23, -0.06]	-0.03	[-0.15, 0.10]
Health status (ref. group: very good)				
good health	3.12	[1.57, 4.83]	0.41	[-1.68, 2.53]
poor/fair health	2.31	[-0.15, 4.63]	1.90	[-1.74, 5.10]
Partnership status (ref. group: single)				
Full-time work	-1.82	[-3.93, 0.24]	-3.20	[-5.63, -0.72]
Part-time work	0.37	[-2.14, 2.94]	-1.30	[-4.28, 1.67]
Not employed	1.82	[-1.24, 4.82]	4.05	[0.81, 7.11]
Self-employed	7.55	[5.58, 9.51]	5.35	[2.81, 7.48]
Location (ref. group: urban)				
Inner regional	2.64	[0.94, 4.48]	1.84	[-1.11, 4.73]
Remote	7.26	[5.27, 9.26]	1.22	[-3.56, 5.76]
Panel B: Men	GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-4.17	[-6.87, -1.47]	-1.70	[-3.60, 0.31]
5-9	-3.10	[-5.62, -0.65]	-1.48	[-3.30, 0.31]
10-15	-2.53	[-4.59, -0.47]	-0.59	[-2.30, 1.05]
Number of children	2.48	[1.19, 3.63]	1.82	[0.92, 2.73]
Age	-0.18	[-0.27, -0.10]	-0.25	[-0.31, -0.17]
Health status (ref. group: very good)				
good health	2.14	[0.61, 3.68]	-0.10	[-1.26, 1.03]
poor/fair health	0.83	[-1.24, 2.79]	0.18	[-1.64, 1.88]
Partnership status (ref. group: single)				
Full-time work	2.39	[-0.23, 5.24]	0.11	[-2.11, 2.47]
Part-time work	0.68	[-2.06, 3.38]	-0.43	[-2.75, 1.79]
Not employed	0.41	[-2.34, 2.97]	-0.57	[-2.80, 1.94]
Self-employed	7.55	[6.10, 8.91]	3.54	[2.37, 4.70]
Location (ref. group: urban)				
Inner regional	1.91	[0.33, 3.43]	-0.50	[-1.99, 0.81]
Remote	4.15	[2.25, 6.02]	-0.20	[-2.78, 2.24]

Table A.V: Marginal effects on hours worked for reduced-form model, imputed wages

Panel A: Women	GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-13.57	[-16.15, -11.00]	-11.64	[-15.19, -8.09]
5-9	-9.69	[-12.15, -7.23]	-6.65	[-10.44, -2.86]
10-15	-4.66	[-6.89, -2.42]	-0.68	[-3.91, 2.54]
Number of children	-0.70	[-1.41, 0.02]	-0.97	[-2.10, 0.17]
Age	-0.16	[-0.25, -0.07]	0.01	[-0.14, 0.16]
Health status (ref. group: very good)				
good health	3.05	[1.32, 4.79]	-0.63	[-3.30, 2.04]
poor/fair health	1.76	[-1.27, 4.79]	-0.32	[-5.16, 4.52]
Partnership status (ref. group: single)				
Full-time work	-0.60	[-3.02, 1.83]	-3.07	[-6.39, 0.25]
Part-time work	1.85	[-1.16, 4.87]	-0.98	[-5.01, 3.05]
Not employed	3.79	[0.92, 6.67]	4.13	[0.92, 7.34]
Self-employed	7.66	[5.56, 9.76]	4.60	[2.35, 6.84]
Location (ref. group: urban)				
Inner regional	1.68	[-0.34, 3.70]	1.88	[-1.46, 5.21]
Remote	6.69	[4.56, 8.83]	1.72	[-2.28, 5.73]
Panel B: Men	GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-5.27	[-8.27, -2.26]	-1.30	[-3.06, 0.45]
5-9	-3.72	[-6.14, -1.31]	-1.55	[-3.19, 0.08]
10-15	-3.06	[-5.13, -0.98]	-1.04	[-2.69, 0.60]
Number of children	1.07	[0.34, 1.81]	0.89	[0.35, 1.43]
Age	-0.35	[-0.44, -0.25]	-0.27	[-0.33, -0.20]
Health status (ref. group: very good)				
good health	2.24	[0.33, 4.14]	-0.31	[-1.83, 1.21]
poor/fair health	0.03	[-2.80, 2.85]	-1.03	[-3.72, 1.66]
Partnership status (ref. group: single)				
Full-time work	7.47	[2.99, 11.95]	2.49	[-0.33, 5.30]
Part-time work	5.74	[1.25, 10.23]	1.40	[-1.39, 4.19]
Not employed	3.94	[-0.57, 8.46]	-0.59	[-3.16, 1.97]
Self-employed	8.80	[7.02, 10.59]	3.69	[2.10, 5.29]
Location (ref. group: urban)				
Inner regional	1.32	[-0.51, 3.14]	-0.39	[-2.17, 1.39]
Remote	4.11	[1.86, 6.35]	-0.99	[-4.67, 2.69]

*Note: Significance is indicated with * for 10% level, ** for 5% level and *** for 1% level.*

Table A.VI: Marginal effects on hours worked for translog benchmark model, observed wages

Panel A: Women	GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-10.84	[-13.08, -7.95]	-10.21	[-12.67, -7.23]
5-9	-8.02	[-10.26, -5.36]	-4.37	[-7.52, -0.16]
10-15	-3.49	[-5.58, -1.31]	0.24	[-2.81, 4.10]
Number of children	-0.56	[-1.17, 0.04]	-0.86	[-1.81, 0.05]
Age	-0.14	[-0.23, -0.04]	0.02	[-0.11, 0.17]
Health status (ref. group: very good)				
good health	2.84	[1.06, 4.58]	-0.12	[-2.16, 2.07]
poor/fair health	1.25	[-1.28, 3.93]	0.82	[-2.46, 4.63]
Partnership status (ref. group: single)				
Full-time work	-4.32	[-6.42, -2.16]	-4.60	[-6.95, -2.16]
Part-time work	-1.23	[-3.98, 1.79]	-2.18	[-5.06, 1.03]
Not employed	3.89	[0.91, 7.09]	4.15	[1.04, 7.17]
Self-employed	7.86	[6.38, 9.86]	4.74	[2.85, 6.72]
Location (ref. group: urban)				
Inner regional	1.78	[0.07, 3.65]	1.16	[-1.63, 4.01]
Remote	6.58	[4.55, 8.85]	1.52	[-3.68, 7.10]
Men	GPs		Specialists	
	Point est.	95% CIs	Point est.	95% CIs
Age of youngest child (ref. group: no dependent children)				
0-4	-3.77	[-6.60, -1.22]	-1.34	[-3.56, 0.67]
5-9	-1.87	[-4.54, 0.49]	-0.99	[-2.86, 0.88]
10-15	-1.77	[-3.97, 0.41]	-0.28	[-2.09, 1.27]
Number of children	1.47	[0.83, 2.13]	1.06	[0.57, 1.51]
Age	-0.16	[-0.24, -0.07]	-0.23	[-0.3, -0.17]
Health status (ref. group: very good)				
good health	1.61	[-0.08, 3.30]	-0.20	[-1.36, 0.91]
poor/fair health	-0.09	[-2.24, 1.92]	-0.19	[-2.12, 1.59]
Partnership status (ref. group: single)				
Full-time work	0.56	[-2.44, 3.47]	-1.26	[-3.23, 0.73]
Part-time work	-1.35	[-4.15, 1.60]	-1.66	[-3.62, 0.36]
Not employed	0.21	[-2.75, 3.30]	-0.33	[-2.40, 1.62]
Self-employed	7.37	[6.12, 8.80]	4.09	[3.19, 5.08]
Location (ref. group: urban)				
Inner regional	1.34	[-0.34, 2.99]	-0.31	[-1.75, 1.04]
Remote	3.91	[1.93, 5.85]	0.17	[-2.02, 2.20]

*Note: Significance is indicated with * for 10% level, ** for 5% level and *** for 1% level.*

Table A.VII: Reduced-form results: OLS of ln(hours)

	Women		Men	
	GPs	Specialists	GPs	Specialists
Ln(hourly wage)	-0.056	-0.070	-0.202***	-0.088***
	(0.132)	(0.059)	(0.070)	(0.030)
Age of youngest child (ref. group: no child or child over 15)				
0-4	-0.444***	-0.337***	-0.121***	-0.029
	(0.043)	(0.052)	(0.035)	(0.020)
5-9	-0.317***	-0.192***	-0.085***	-0.034*
	(0.041)	(0.056)	(0.028)	(0.018)
10-15	-0.152***	-0.020	-0.070***	-0.023
	(0.037)	(0.048)	(0.024)	(0.018)
Number of children	-0.023*	-0.028*	0.025***	0.019***
	(0.012)	(0.017)	(0.009)	(0.006)
Age	0.016	0.062***	0.056***	0.083***
	(0.012)	(0.017)	(0.010)	(0.011)
Age squared	-0.023*	-0.067***	-0.063***	-0.087***
	(0.013)	(0.017)	(0.010)	(0.011)
Health status (reference group: very good)				
Good health	0.100***	-0.018	0.051**	-0.007
	(0.029)	(0.039)	(0.022)	(0.017)
poor/fair health	0.058	-0.009	0.001	-0.022
	(0.051)	(0.071)	(0.033)	(0.030)
Partner's employment (reference group: single)				
not employed	0.124***	0.119**	0.091*	-0.013
	(0.048)	(0.047)	(0.053)	(0.029)
works full-time	-0.020	-0.089*	0.171***	0.054*
	(0.040)	(0.049)	(0.053)	(0.031)
works part-time	0.061	-0.028	0.132**	0.031
	(0.050)	(0.060)	(0.053)	(0.031)
Self-employed	0.251***	0.133***	0.202***	0.081***
	(0.035)	(0.033)	(0.021)	(0.018)
Location (reference group: urban)				
Inner regional	0.055	0.054	0.030	-0.009
	(0.034)	(0.049)	(0.021)	(0.020)
Remote	0.219***	0.050	0.094***	-0.022
	(0.035)	(0.059)	(0.026)	(0.041)
Other income	-0.008***	-0.001	-0.008***	-0.007***
	(0.003)	(0.003)	(0.002)	(0.002)
Constant	3.513***	2.633***	3.303***	2.373***
	(0.628)	(0.487)	(0.362)	(0.293)
Number of observations	1067	769	1128	1908
Adj. R-squared	0.2885	0.1794	0.2887	0.2368

*Note: Significance is indicated with * for 10% level, ** for 5% level and *** for 1% level.*

References

Ribeiro, E. P. (2001). Asymmetric labor supply. *Empirical Economics 26*(1), 183–197.

Van Soest, A. (1995). Structural models of family labor supply: a discrete choice approach. *Journal of Human Resources*, 63–88.