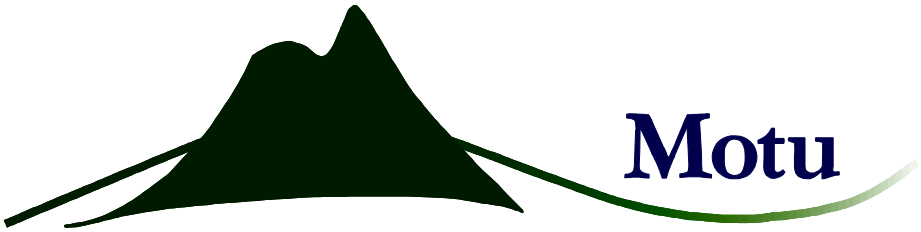


Addressing the absence of hours information in linked employer-employee data

Richard Fabling and David C. Maré

Motu Working Paper 15-17
Motu Economic and Public Policy Research

October 2015

Author contact details

Richard Fabling
Independent Researcher
richard.fabling@xtra.co.nz

David C. Maré
Motu Economic and Public Policy Research & The University of Waikato
dave.mare@motu.org.nz

Acknowledgements

The techniques used in this paper have been developed and refined over a number of years with valuable input from co-authors, particularly Dean Hyslop and Lynda Sanderson. We thank Productivity Hub agencies for funding this work, and Statistics New Zealand for supplying and enabling access to the data.

Disclaimer

The results in this paper are not official statistics, they have been created for research purposes from the Integrated Data Infrastructure (IDI) managed by Statistics New Zealand. The opinions, findings, recommendations and conclusions expressed in this paper are those of the authors not Statistics NZ, Motu, or The University of Waikato.

Access to the anonymised data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business or organisation and the results in this paper have been confidentialised to protect these groups from identification.

Careful consideration has been given to the privacy, security and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the privacy impact assessment for the IDI available from www.stats.govt.nz.

The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. These tax data must be used only for statistical purposes, and no individual information may be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes. Any person who has had access to the unit-record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to secrecy. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

Motu Economic and Public Policy Research

PO Box 24390
Wellington
New Zealand

Email info@motu.org.nz
Telephone +64 4 9394250
Website www.motu.org.nz

ISSN 1176-2667 (Print), ISSN 1177-9047 (Online).

Abstract
The availability of tax-based payroll data has proved a blessing to labour and business economists wishing to understand workers, their jobs and their employers. Unfortunately, administrative data do not always include key variables of interest. In the case of New Zealand, linked employer-employee data do not include any information on hours worked. We implement a set of complementary methods to patch this gap, deriving an approximate measure of full-time equivalent labour input. In addition, and more specific to the New Zealand data environment, we describe a method for identifying working proprietors using annual tax-filed information, thus providing a more complete picture of total firm labour input.

JEL codes
D22; L11; Q54

Keywords
Linked employer-employee tax data; measuring labour input; full-time equivalent; working proprietors

1 Motivation

Being able to accurately measure labour input (L) is important to the study of individual labour market outcomes (such as the returns to education), and to the analysis of firm performance (for example, productivity estimation). New Zealand is blessed with administrative tax data covering all paid jobs, allowing accurate identification of employing businesses, total gross earnings and monthly counts of employees.

However, these data miss a vital component – hours worked – which, if available, would substantially improve the measurement of L. In the absence of hours information, workers are often assumed to be all be full-time resulting in wage rates being underestimated for part-time workers, and productivity being underestimated in businesses that make relatively more use of part-timers. The degree of this mismeasurement is potentially large, since 28% of employees are part-time (on average, working less than 30 hours a week) according to the latest official Household Labour Force Survey (HLFS) statistics. The importance of hours measurement is even clearer when we consider analysis involving gender differences in earnings, since there is a substantial gender gap in part-time status between men (16%) and women (40%) (Statistics New Zealand 2015).

In response to this hours data gap, we propose a method of deriving an approximation to individual monthly labour input using a plausible set of assumptions about workers, namely that: the statutory minimum wage is observed; multiple-job workers have the same total labour supply as single-job workers; and that hourly wage rates are likely to be constant over adjacent months (at the start and end of jobs). As part of this derivation, we also identify job spells correcting for short (one month) breaks in earnings, which may be useful to researchers analysing job transitions.[1]

While the resulting adjusted labour input will still overestimate true L for a subset of workers, we believe the method is materially superior to a simple headcount approach. At a minimum, the method identifies – through a series of indicator variables – a subset of job months that researchers may wish to exclude from any employment analyses that assume constant labour input across worker, at least as a robustness check on their main results.

To complete the picture of labour input at the firm level, we also introduce a new definition of working proprietors (WP) and their associated

[1]Identifying job spells and adjusting job starts and ends are the main innovations that we introduce to the precursor work of Maré and Hyslop (2006).

annual labour input. This identification ensures that a large population of firms with measured output but no employees (predominantly sole proprietorships and partnerships) have positive measured total labour input. Because of the nature of the self-employed income data, this method overestimates WP labour input, as it yields an annual headcount and inevitably counts some passive investors as if they were working proprietors. While we attempt to limit the inclusion of passive investors, these measurement issues suggest caution when analysing, for example, the productivity of WP-only (and WP-dominant) firms (Fabling and Sanderson 2014).

A key purpose of the WP identification is to exclude these individuals from the derived measure of employee labour input. The separation of WPs from employees in the wage data removes individuals who determine their own "wage," and who may receive supplemental income through non-wage channels (eg, paid as dividends). Stripping these individuals from the employee data may be an important step in the accurate assessment of worker outcomes, depending on the population of interest, since WPs with "wage" income from their own business constitute approximately two and a half percent of employee jobs.[2] Additionally, over a quarter of working proprietors also hold a job as an employee in a different firm, suggesting that identifying their income from business ownership may be a necessary step to understanding their concurrent behaviour as employees of other firms.

Section 2 outlines the available administrative data and the implications that this has for the employer unit of analysis and for the temporal dimension of each of the components of L. Sections 3 and 4 cover the derivation of WP and employee labour input, respectively. Section 5 illustrates, using three examples, the impact of FTE adjustment on derived results, as compared to a simple headcount measure of labour input. Section 6 provides a summary. Researchers who wish to use the derived tables but who do not need a detailed understanding of the methodology can skip directly to appendix A, where we present data dictionaries for all the derived tables for ease of reference. Appendix A also documents the three differences between the final methodology documented in this paper and the methodology used to derive the current suite of tables available in the research database.[3]

[2]Twelve percent of working proprietors receive EMS income at some point from a business that they own.

[3]These differences are also noted as footnotes in the main text.

2 Data, statistical units and timeliness

2.1 Employees and jobs

The derivation of an approximate full-time equivalent (FTE) labour input for employees relies on two key inputs, both located in Statistics New Zealand's Integrated Data Infrastructure (IDI) – the Employer Monthly Schedule (EMS) from Inland Revenue, and the Business Register (BR) constructed and maintained by Statistics New Zealand.

The EMS is filed by all businesses with paid employees, since employers are obliged to deduct Pay-As-You-Earn (PAYE) income tax at source. These data sit within the `ir_clean` schema of the IDI in the `ird_ems` table. The raw data relate two confidentialised IR numbers (a payer and a payee), therefore jobs are identified at the tax filing unit level, which is almost always the enterprise level. Confidentialised payer IR numbers are mapped to enterprise numbers by Statistics NZ, using relationships held on the BR.[4] These, in turn, are mapped to permanent enterprise numbers (PENTs), which are enhanced longitudinal business identifier that use employee-tracking to repair broken enterprise number linkages (Fabling 2011). Similarly, Statistics NZ map confidentialised payee IR numbers to worker ids (`snz_uids`) which, in turn, link to other person-level collections in the IDI. Thus, ultimately, a job is defined as a worker-firm (`snz_uid` $\times$ `pent`) relationship observed at a monthly frequency.

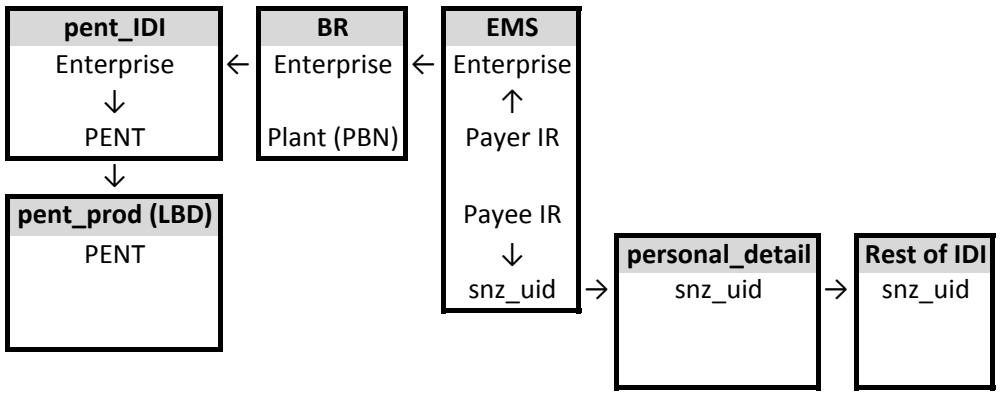
Statistics NZ allocates workers to the physical locations (plants) of the business. This is done within the Linked Employer-Employee Dataset (LEED) methodology and carried over to the EMS table in the IDI. Plants are identified by Permanent Business Numbers (PBNs) For businesses with multiple active locations on the BR, this allocation is done by balancing the expectation that workers live within commuting distance of their job against a desire to mimic known (or approximate) employment shares across the locations.[5] The accuracy of this allocation is hindered somewhat by the presence of low quality residential address information for some employees (ie, out-of-date or postal, rather than residential, addresses).

As part of the design of the LEED methodology, workers are usually

[4] We lose a small number of observations that have untraceable payee (worker) IR numbers, or which are associated with businesses that have not been created ("birthed") on the BR.

[5] Known plant-level employment shares largely come from the Quarterly Employment Survey and the Business Register Update Survey, both administered by Statistics NZ.

Figure 1: Schematic of employer & employee id links across key tables



In practice, the confidentialised payer and payee IR numbers are labelled `snz_employer_ird_uid` and `snz_ird_uid` respectively in the `ird_ems` table.

kept at their initially allocated location over time but may be reallocated, meaning that any apparent worker movement between two plants in a firm is unlikely to represent the timing of real-world events (ie, internal job transfer within a business). For this reason, we do not use these changes to infer anything about jobs or job spells. We do, however, retain Statistics NZ's mapping of workers to PBNs for end users.

Finally, since minimum wage rates were historically age-dependent, we make use of the recorded birth date on the `data.personal_detail` table. This table records the "best" guess, collated across all the relevant data sources in the IDI, of birth date (month and year) and gender of the individual. We retain both age and gender in the final job-month table for convenience. Figure 1 presents a schematic of this data set-up for reference.

2.2 Working proprietors

While some working proprietors can be identified entirely from within the EMS data – ie, when the payer and payee have the same IR number – most WP labour input is determined using additional tax filings. These returns relate to annual income declarations for individuals (IR3) and partnerships (IR7), together with company shareholder details (IR4S). Forms, including

the EMS (IR348), are included in appendix B.[6] Together these data identify income derived from self-employment in business.

Annual tax return data relate directly to firm-level variables (eg, total partnership income) and, at present, access to such business IR data requires researchers to be government employees. In the current IDI, non-government researchers have the ability to see worker-level information on monthly earnings, but not the employer identifier, so that it is not possible to determine the aggregate wage bill or other properties of the employer. These same access constraints prevent non-government researchers from accessing the LBD, which relies heavily on tax-based business financial information (Fabling 2009; Fabling and Sanderson 2015), and includes the aggregate firm-level measures of L we derive (described in appendix A).

Legally, business owners do not necessarily have to be people (eg, partnerships may be part-owned by companies). Therefore, we additionally need to determine which IR numbers are and aren't associated with individuals. As a first step, we make use of the fact that confidentialised payee IR numbers in the EMS must belong to individuals. To identify confidentialised IR numbers associated with individuals who have never received PAYE income, we then rely on IR's Customers table, `ir_clean.ird_customers`, which identifies confidentialised IR numbers of individuals (`ir_cus_entity_type_code`=I).

Finally, since WP information is largely annual, but includes a monthly dimension introduced by the potential for payment via the EMS, we also need to know business balance date information to enable the monthly data to be aggregated to the appropriate tax year.[7] Balance date information comes from the Business Register, and is populated from information supplied by IR to Statistics NZ.[8]

[6] Forms relate to the 2012 tax year, but are representative of all years of data included in the IDI. Historically, partnership distribution of income was filed as part of the IR7, but is now separately reported in the IR7P (coinciding with the coverage of the IR7 expanding in 2012 to include reporting from look-through companies). We will refer to these data collectively as the IR7. Most variables collected in these forms are not currently available in the IDI.

[7] Strictly speaking, balance dates relate to financial years, but IR strongly recommends that businesses align their balance date to their mandated tax year, which for most firms is a 31st March year-end. In the LBD, this common year is labelled the `dim_year_key` and is represented by the closing year followed by "03" – eg, 201103 for the tax year ending 31 March 2011. Businesses with non-March balance dates have their returns assigned to the representative March-based `dim_year_key` that most overlaps with their financial year.

[8] Five percent of firms with positive total L have multiple recorded balance dates on the Business Register (eg, because they applied to IR for a balance date change). These firms are assigned a permanent (primarily employment-based) predominant balance date

2.3 Data timeliness

The IDI is updated quarterly and results presented in this paper relate to the December 2014 instance, labelled `IDI_Clean_20141205`. The time stamp in the database name refers to the build completion date (5 December 2014), rather than representing the latest period covered by the tables. Coverage periods are dataset-specific and depend on the timeliness of the contributing administrative collections, and the frequency and timing of when each collection is supplied to Statistics NZ for integration into the IDI.

Employers are required to file EMS returns to IR by the 20th of the month following the payment month, and the majority of these returns are filed electronically. Statistics NZ receives a monthly feed of these data from IR. In practice, this results in the latest (approximately) complete EMS month available in `IDI_Clean_20141205` being March 2015, which is roughly nine months behind the build completion date.[9] Since these dates are typical of the update process and IDI updating is quarterly, the latest complete EMS month is generally at most twelve months behind the current date. EMS data start in April 1999, meaning that 15 complete years (ending in March) are currently available.

Annual filed tax information also extends back to the March 2000 year, but the most recent data are less timely than the EMS because their availability relies on the completion of financial accounts after the close of each tax year. For example, for an individual (sole proprietor) filing an IR3, that return must be filed by the 7th of July following the (March 31st) year end.[10] Filers who use tax agents gain extensions beyond this date to smooth the workload of the agents. As a result of the use of agents some returns are legitimately filed as late as the following March.[11] These same filing patterns apply to tax data used in the productivity dataset (Fabling and Maré 2015), as the necessary financial (IR10) information is submitted by many businesses at the same time as their income tax returns, providing supplemental evidence of declared taxable profit.

---

to ensure that all `dim_year_keys` contain 12 months.

[9] There are a further three months of EMS data present in this instance of the IDI (ie, April 2015-June 2015), but coverage tails off substantially over this period and these data haven't been used due to the risk of over-identifying job spell ends in the tail of the data.

[10] For balance dates between 1 October and 31 March (inclusive), the due date for filing income tax returns is 7 July. For balance dates between 1 April and 30 September (inclusive), the due date is the seventh day of the fourth month after the balance date.

[11] Currently IR sets targets for tax agents of 40% of returns completed by end September; 60% by end November; 80% by end February and 100% by the following 31st March.

As is the case with the EMS data, the quarterly updating of the IDI means that annual tax data are included as they are supplied by IR. In general, this results in partial coverage data being available for the most recently completed tax year. We use these data to identify WP-firm relationships but cannot construct a complete measure of WP labour input in that year. In the current instance of the IDI this means that the employee measure of L is available for one more (March) year than the WP data, and that some WPs may be incorrectly identified as employees, particularly in firms whose first year of operation is the year without complete annual tax filing data (currently 2014).

3 Identifying working proprietors

In order to eliminate business owners from the subsequent employee FTE calculation, we calculate WP labour input first.[12] Individuals are identified as working proprietors of a business if they receive any of the following types of income from the business (relevant tax forms are included as appendix B for ease of reference):[13]

1. *Sole proprietors paying themselves PAYE income,* defined as EMS payments where the payer and payee IR numbers are the same
2. *Sole proprietors receiving non-zero self-employment income* as reported in box 23 of the IR3 (variable `ir_ir3_net_profit_amt` in table `ir_clean.ird_rtns_keypoints_ir3`)
3. *Partners receiving a share of total partnership income* as reported in box 25J of the IR7P (variable `ir_ir20_tot_share_of_inc_865_amt` in table `ir_clean.ird_attachments_ir20`)[14]

[12] We use the terms business owner and working proprietor interchangeably even though the latter is a subset of the former. This is simply for convenience. In practice, we do not have the data to identify the business owners comprehensively and, where we do have data, we actively exclude passive investors who would count as business owners but not working proprietors.

[13] In addition, the payer IR number must be coded to an enterprise number on the BR to establish this relationship. Otherwise, while there may still be a WP-business relationship, there are no business characteristics available for research purposes and the business will not have a PENT identifier.

[14] This table is labelled "IR20," but a search of IR's website does not produce any matches to this phrase. We refer to the data as IR7 filing throughout the paper, consistent with the form that currently collects the information.

4. *Company owners receiving remuneration with no PAYE deducted* as reported in box 41C of the IR4S (variable `ir_ir4_tot_sholder_sal_809_amt` in table `ir_clean.ird_attachments_ir4s`)

We assume that these ownership relationships are permanent (time-invariant), so that any individual ever observed to be a working proprietor is deemed to be one over the life of the business.[15] This rule is not equivalent to assuming WPs supply labour in all years – WPs are counted as working only in years where they receive profits or PAYE income from the business. In practice, therefore, the permanent view of ownership status affects only the treatment of PAYE income accruing to WPs. This approach adds stability to the measurement of both WP and employee labour input over time, which may be beneficial if WPs vary their method of paying themselves. For example, if a WP received a base salary via the PAYE system and then received additional income (via annual filings) only in years where the business performed well, then adopting a time-varying view of WP status would, in this hypothetical example, see the individual counted as an employee in bad business years and as a WP in good years.

The potential downside of assuming permanent relationships is neglecting the possibility that transitions from EMS income receipt to annual self-employed income (and vice versa) may identify real-world changes in business ownership status. Mitigating this concern is the fact that 88 percent of WPs never receive EMS income from the business they own. In addition, of those who receive both EMS and annual forms of income there is a tendency for EMS income to appear *within* the timespan over which annual income is earned. Specifically, two thirds of WPs do not receive any EMS income from the business prior to the year of their first annual income, and just over half do not receive EMS income in years after they stop receiving annual income. Including EMS payments that pre-date (postdate) annual payments by at most a year, the rate at which filing patterns for EMS are "contained within" annual spells rises to 78 (68) percent.[16]

The IR4S (company) declaration rule potentially overcounts WPs, since this form of remuneration is paid to "shareholders, directors and relatives of shareholders" (appendix B). To limit the extent of this overcounting, we

[15] This method is inconsistent with the LEED measure of self-employed income, which uses a mix of annual observations for some data sources and permanent for others.

[16] These statistics exclude one-off payments of either income type to avoid comparing single period "income spells." To accurately define the first (last) payment year, the initial first (final last) year is restricted to 2004 (2009). That is, for example, the first year of EMS is preceded by at least four years of non-EMS payment.

apply a minimum earnings threshold, which in an attempt to eliminate non-WP payees. The threshold is set at earning over $15,000 (in real 2000 dollars, deflated using the CPI), determined with reference to the distribution of actual payments.[17]

Figure 2 (panel A) shows the distribution of IR4S income by five thousand (real, 2000 year) dollar bands. The first three columns above zero capture payments below the threshold value and, ignoring the shading of the bars for the moment, demonstrate the relatively large drop-off in incidence of payments to the right of the cut-off, which was the determining factor in setting the cut-off level.[18]

A significant proportion (41 percent) of IR4S payments lie to the left of the assumed cut-off, implying that this choice is material to the calculation of company WP labour input. Having said that, when implementing the cut-off, we require only that IR4S income exceed the threshold value in at least one year. If this condition is met, we count as WP all years where the owner received IR4S income regardless of the quantum, including those years that precede the first year where income is above the inflation-adjusted threshold. The effect of this secondary choice can be seen from the shaded (included) observations in the income distribution that sit to the left of the cut-off.[19] Overall, this approach results in 46 percent of observations below the cut-off being excluded, which is equivalent to 19 percent of all observations.

This "permeable barrier" approach avoids situations, for example, where an individual is consistently paid an amount near the cut-off, but is included as a WP in only a subset of years because of the deflator choice. It is also consistent with the other forms of WP identification, which do not depend on a threshold income level. Panels B and C of Figure 2 show the comparable data for sole proprietors and partnerships respectively, including those returns associated with negative profit distribution. As they should, these

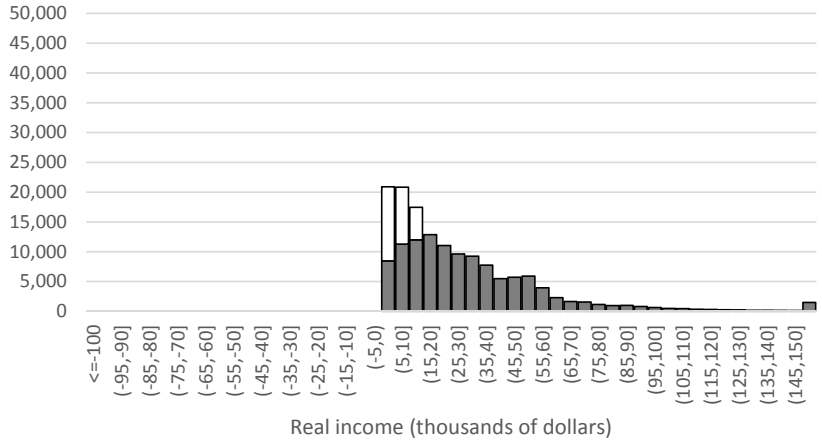
[17] LEED takes the alternative approach of making use of the Inland Revenue cross-reference table, which provides information on the relationship between two IR numbers. In principle this approach is superior, as the table theoretically allows the user to separately identify directors and other officeholders from business owners. However, in practice the data in the cross-reference table are often inconsistent with filed returns, suggesting that it is not a particularly reliable source for this purpose.

[18] The actual cut-off assessment was made using narrower band widths and the same pattern is still apparent. All years have been pooled and counts are annualised. All counts in this paper have been random-rounded in compliance with Statistics NZ confidentiality rules.

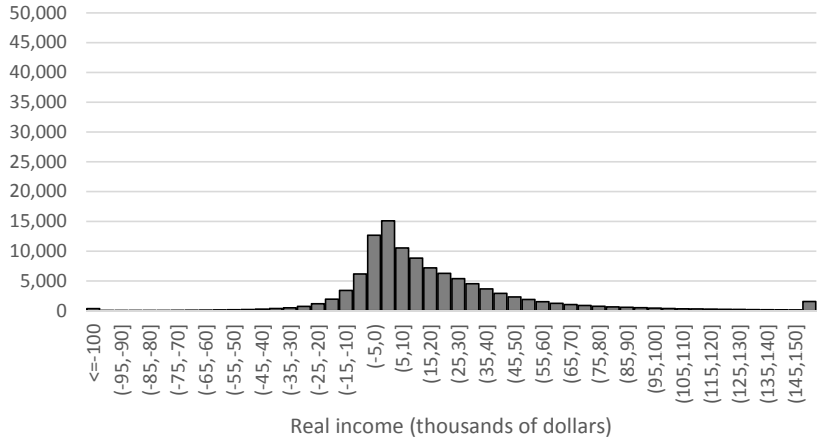
[19] Derived tables based on the December 2014 instance of the data do not implement this secondary rule and count (non-W&S based) company WPs only when IR4S income in the year exceeds the (real) threshold value. December 2014-specific rules are discussed further in appendix A.3.

Figure 2: Working proprietor count by real income distribution

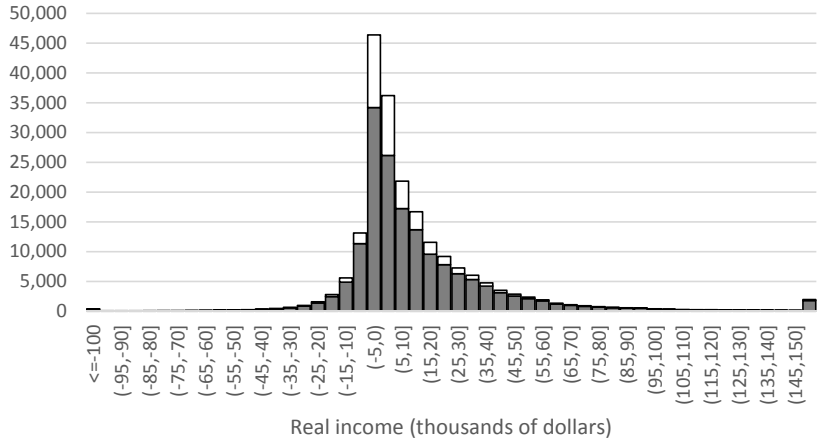
A. Companies (IR4S)



B. Sole proprietorships (IR3)



C. Partnerships (IR7)



All years (2000-2013) pooled, income deflated using the average CPI (2000 dollars), and counts annualised. Returns are excluded if the payer is not on the Business Register and/or the payee is not an individual. Shaded area counts individuals included as working proprietors. Unshaded area is individuals not meeting the minimum income threshold (panel A) or with zero IR3 partnership income (panel C). Negative (positive) profit distribution below -$100K (above $150K) pooled.

income profiles bear close resemblance to measures of profitability derived at the firm level, which display density clustered around moderate positive rates of return and a clear asymmetry between the left and right tails with far fewer negative than positive profit firms. The presence of negative income (profit distribution) at the WP level – representing, on average, 26 (35) percent of observations for sole proprietors (partners) – clearly reinforces the principle that observed payments are not purely related to labour compensation. This factor justifies the need to ignore the quantum of income when identifying active sole proprietors and partners since, for example, an individual WP may supply the same labour input over adjacent years, but receive markedly different self-employed income across those two years.

Inclusion of profit distribution also means that WP counts may be overestimated because payments purely reflect a return to capital invested, rather than including a payment for labour input. This is a particular problem for partnership income, which constitute 45 percent of self-employed income observations, since recorded returns from these businesses are based on agreed profit shares and not necessarily linked to labour input. For this ownership type, though, we apply an additional test to assess whether the partner was a passive investor. We can do this because partners who gain their partnership income purely through capital investment are required to report that income in box 24 (other income) of their individually-filed IR3, rather than box 18B, which records active partnership income (see appendix B).[20] Therefore, the definition of partner (WP) excludes those individuals who receive partnership income on the IR7, but who didn't also declare positive partnership income on the IR3. Excluded observations under the additional partnership rule are represented by the unshaded component of the bars in figure 2 (panel C). As the figure shows, most of these observations (79 percent) fall within plus or minus $15,000, and 20 percent of all IR7 observations are dropped because of the secondary (IR3-related) test.

Other identifying forms of WP income seem more likely to be associated with direct labour input. In all instances of self-employed income declaration though, reported total remuneration to WPs may contain a return-to-capital component. It is for this reason that we do not attempt to assess the magnitude of WP input based on reported income, nor the relative contribution of WP labour for those individuals who also work as an employee for another employer during the year.[21] Instead, we simply count WP input as equiva-

[20] This distinction is made because active partners in a business pay Accident Compensation Corporation (ACC) levies associated with their labour, whereas passive partners do not.

[21] Any attempt to coordinate the allocation of labour between WPs and other jobs would be further complicated by the frequency difference (ie, annual versus monthly), making

lent to that of a full-time full-year employee. We then impose that total WP labour within a year sums to one so that individuals who own and work in multiple businesses have their measured input distributed evenly across all the firms they own (ie, WP labour input is the reciprocal of the number of firms owned).

Since we do not attempt to compare WP and employee labour input, a subset of individuals' will have total labour input greater than one when measured across WP and employee jobs. The proportion of WPs who own multiple businesses and/or hold a job as an employee of another business is shown in table 1. On average, two thirds of WPs own a single business and do not hold a job as an employee for another business. A further 26 percent own and work in one business whilst holding a paid wage and salary job in another business and, therefore, may have a measured total labour input above one. The remaining 7 percent of WPs own and work in multiple businesses during the year, and approximately one in five of those individuals (1.4 percent of total WPs) also has paid employment elsewhere. Given the substantial rate of business turnover for small business, multi-WP counts potentially include many individuals winding up one business venture and starting another during the same financial year. Consistent with this view, multi-firm ownership is more likely to be associated with a first/last year of WP status.[22] Similarly, WPs may appear to simultaneously hold other jobs because of transitions to and from self-employment.

Because of this logic, the final step in the WP calculation imposes a cap of 0.5 WP labour input in "transition years", defined as years where the WP was not observed at the firm in the prior and/or following year.[23] These transitions are a common occurrence in the WP data. For example, over the 2000-2008 years an average of 16.5 percent of WP observations are transition years based on inactivity in the following year (transition to exit). If we further break that number down into "permanent" and "transitory" exit based on whether the WP is ever observed subsequently at the firm after the break, permanent exit affects 11 percent of observations, with transitory exit making up the other 5.5 percent. Of the transitory exits, approximately half

it difficult in some instances to establish whether roles as business owners and employees elsewhere are held concurrently or sequentially.

[22] Specifically, 27 percent of multi-WP observations are associated with the first and/or last year of being a WP at the firm. For single-WP individuals, the comparable statistics is 20 percent.

[23] The current (December 2014) instance of the tables do not implement this transition adjustment for WPs. Appendix A.3 explains the flow-on effects of this difference in terms of how the data are used for productivity measurement.

of the breaks before the WP receives income from the business again are at least two years long.

The WP transition adjustment recognises that labour input is expected, on average, to be lower in years adjacent to breaks in activity.[24] Because WPs with multiple firms have L set to the reciprocal of the number of firms owned, the transition cap has no effect on their measured labour input. Table 2 summarises the proportion of transitions by year, broken down by whether the transition is an entry or exit. Over a quarter (25-29 percent) of observations are transitions, meaning that 21-23 percent of single firm WP L is set to 0.5 (right-hand column). While this rate of adjustment might seem high, the statistics on exit suggest that two thirds of these WP transitions are permanent – that is, that the high turnover rate of WPs at firms is driven by dynamic entry and exit of micro enterprises, rather than transitory reporting of income associated with ongoing businesses. Some of this apparent dynamism in "permanent" entry and exit may, in turn, be due to the inability of the PENT technology to repair enterprise number breaks for enterprises without employees.[25]

It is difficult to identify transitions in the first and last complete year of WP data. In these years we set transition flags to null where potential transitions are unobservable. However, for the last complete year, there are a subset of observations where we can confirm *non*-transitions. Using the December 2014 IDI as an example, partial identification of non-exit following the 2013 year come from complete EMS filing for 2014 together with incomplete annual tax data for that year.[26,27] Because partnership and company

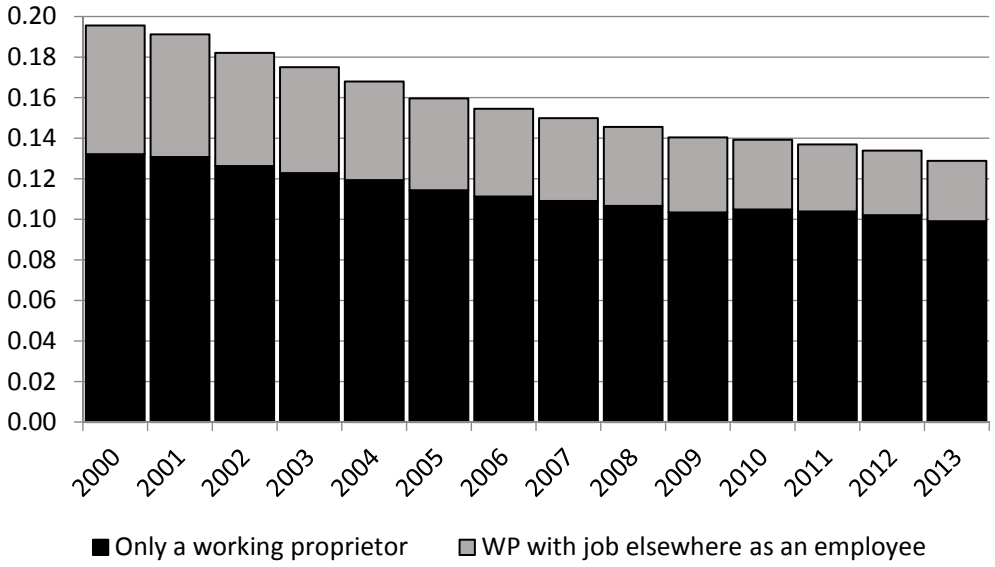
[24] In the case of WP-only (non-employing) firms, such transitions may also be associated with transitions in and out of *any* business activity.

[25] Fabling (2009) demonstrates that one effect of the employee-tracking PENT technology is to enable identification of businesses changing legal form from sole proprietorships and partnerships to limited liability companies. Such transitions may potentially be important to the WP-only sub-population of businesses, inflating the apparent entry and exit rate of WP-PENT relationships when such transitions cannot be repaired. Given the availability of WP identifiers it may be possible to repair WP-only enterprise number links using these ids in an analogous way to the employing firm repairs, perhaps also using information on business location and activity to triangulate continuity.

[26] Annual tax data for 1999 are not available. While some firms have balance dates that mean that the EMS data covers part of the 199903 `dim_year_key` this is a small and selective group (because balance dates relate to industry characteristics), so the same methodology is not applied to the first year of data.

[27] Derived tables based on the December 2014 instance of the data do not implement this transition adjustment. The number of individuals who are identified as WPs using the incomplete 2014 year is 34 percent of the 2013 count of WPs, suggesting that these rules can materially improve the identification of transition-to-exit in the final year of complete

Figure 3: Share of individuals who are working proprietors



Proportion of distinct individuals in the final dataset who are working proprietors during the year. Grey shading shows those WPs who are also employees in a job outside of the owned firm, calculated from the EMS data assuming a common 31st March year end.

filing is done at the firm-level, and because the 2014 EMS year is complete, we can also confirm WPs in 2013 that will exit in the following year for WPs at these business types where the business has filed and the WP is not in receipt of income according to the annual return and the EMS. Similarly for sole proprietors, we can identify those WPs who have filed IR3s for the 2014 year, but have reported zero self-employed income (and have not paid themselves an EMS wage).

Multiple firm WPs, WPs in 2000 who will exit in 2001, and WPs who entered in 2013 also have known L because of the choice of a fixed cap transition adjustment. The former group is unaffected by the cap, and the other two groups are subject to the same transition year adjustment regardless of the unknown state on the other side of the current year. Remaining observations of L are left unadjusted (ie, set equal to one) and users should exercise caution in situations where these data are needed. Appendix A.2 describes how Fabling and Maré (2015) exploit firm-level data to adjust final year observations for micro productivity measurement purposes.

data. December 2014-specific rules are discussed further in appendix A.3.

Overall, working proprietors make up between 13 and 20 percent of all individuals in the labour data (ie, WPs and employees), depending on the (March) year. Figure 3 shows how this proportion changes over time, broken down by whether the working proprietor has a job elsewhere as an employee during the year. The proportion of individuals who are working proprietors has declined substantially over time – both for those who are only WPs (falling from 13.2 in 2000 to 9.9 percent in 2013), and for those who additionally have a job as an employee in another business (falling from 6.3 to 3.0 percent).

4 Calculating FTE labour input

The calculation for employees exploits the principle that workers, like WPs, have a common fixed maximum amount of labour to allocate in any month, normalised to a unit of one (a full-time equivalent, FTE). The method proceeds in three primary steps. Firstly, we identify a subset of the workers whose labour input – across all jobs worked – is likely to be less than full time, deriving an estimate of total labour input for them. Secondly, we allocate total L between jobs for multiple job holders. We make this allocation using income shares. Finally, we identify job start and end months and adjust labour input to account for the fact that jobs tend to start and end with part-months. We explain each of these steps in turn.

4.1 Approximation of total monthly labour input

By looking at aggregate monthly earnings from wages we identify workers who are likely to be less than full-time, because the worker has total gross earnings from wage & salary below the amount a minimum wage worker would earn in a month, assuming a 40 hour work week (and 4.35 weeks in an average month).[28] The total (normalised) L of a worker (i) across all jobs is calculated as

$$L_{it}^{tot} = \min(1, \frac{E_{it}^{tot}}{40 \times 4.35 \times m_{it}}) \tag{1}$$

where E^{tot} is the sum of earnings across all jobs in the month (t), m is the applicable minimum wage rate. Minimum wage rates are shown in table

[28]This is the only part of the calculation that sets any expectation regarding the number of hours that a "full-time equivalent" works.

3 and, in some years, are age-dependent. For workers above the minimum wage, their total FTE is unadjusted by this stage.

4.2 Allocation between jobs for multiple job holders

At the second step, L_{it}^{tot} is allocated between all the jobs (j) held by the worker during the month. For multiple job holders, this allocation is done using relative earnings in the month, specifically

$$L_{ijt} = L_{it}^{tot} \times \frac{E_{ijt}}{E_{it}^{tot}}. \tag{2}$$

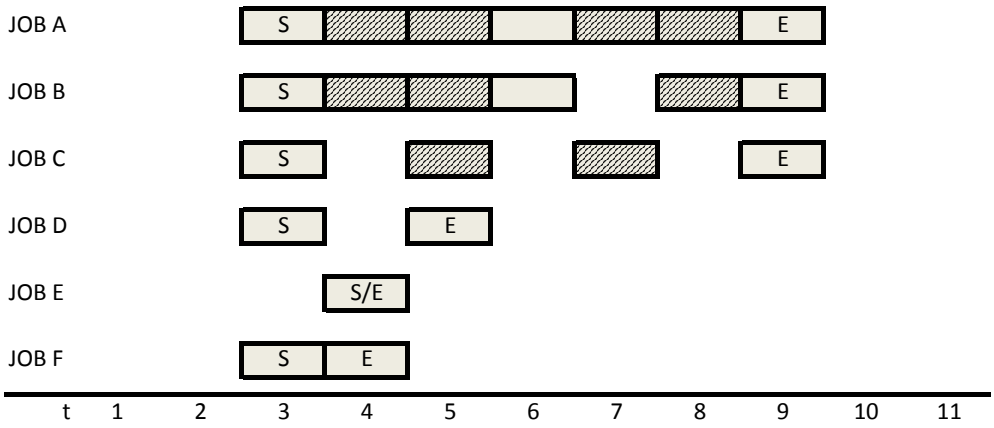
This allocation mechanism will be accurate if a worker receives the same hourly rate across concurrently held jobs. By construction, multiple job allocation applies to single-job workers changing employer, if the job change results in payments from both the old and new employer occurring in one (or more) transitional months. This particular case may represent a scenario where the common hourly rate assumption is violated, since workers tend to move to higher paying jobs (see, eg, Maré et al. (2014) for evidence of this in New Zealand), introducing a potential bias into the calculation – ie, overallocation of L to the new (starting) job. Conversely, if the ending job triggers an exit payment, then the bias may be in the opposite direction, so that labour is overallocated to the ending job. The final step in the FTE derivation partially corrects this bias, at least for one of the jobs.[29] If the worker takes sufficient time off between jobs it may be possible to adjust both start and end L to be better approximations of their true values than is generated by income share allocation.

4.3 Adjustments for job starts and ends

We expect that most workers will work a part-month when they start or end jobs. As a consequence, the measured FTE should, on average, be lower in these months, compared to adjacent interior months. In order to identify start and end months, we first need to define what we mean by a job spell. We then use up to two adjacent interior months of wage information for

[29] For example, assuming no exit payments, and since the hourly rate in the new job is, on average, higher than the old job, the FTE allocation to the new job is too high and, conversely, L_{ijt} for the old job is too low (assuming the employee does not take time off between jobs). Because each step of the FTE derivation only revises FTE estimates downwards, only the new job can be further adjusted to the correct level.

Figure 4: Examples of job spells



Job A-C represent single employment spells from $t = 3$ to $t = 9$. Darker shaded periods indicate months used to calculate interior wage rates for start (S, $t \in \{4, 5\}$) and end (E, $t \in \{7, 8\}$) month adjustment. Jobs D-F illustrate the three distinct types of short spell possible when job spells (spells without interior months) are allowed to have one month gaps.

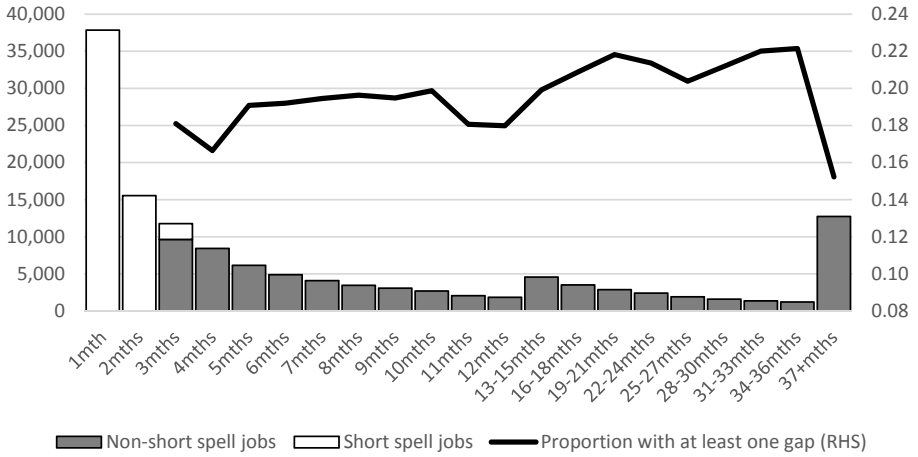
the same job, where available, to adjust start and end month FTE on the assumption that the wage rate is likely to be constant over short periods of time.

Job spells are defined to be periods of employment with the same employer and with at most one month breaks in observed earnings. While we allow only one month breaks, a worker may have multiple such breaks during the job spell. Figure 4 illustrates this definition, where each rectangle represents a monthly EMS payment. Jobs A-C all show jobs spells starting at time $t = 3$ (labelled "S") and ending at time $t = 9$ (labelled "E"). These are deemed to be the start (end) dates because there are no earnings in the previous (next) two months, and there are no earnings breaks longer than a single month between these two dates.

The first job, job A, is the simplest case where the worker has been paid in each month of the employment spell, whereas job B has the simplest short break, a single month at time $t = 7$. EMS payments in job C are less frequent, but still conform to the simple rule of having at most one month earnings breaks during the job spell.

Darker shaded cells in figure 4 represent the interior months that are used to infer hourly wage rates in the start and end month. Jobs B and C show why we use the two months after (before) the spell start (end). Be-

Figure 5: Number of spells starts by duration, and prevalence of gaps



Counts are the monthly average number of jobs of a given duration and starting over the period June 1999 to January 2011. Short spells (spells with no interior months) are unshaded. Spell durations greater than 12 months are grouped into calendar quarters and durations above 36 months are pooled.

cause the spell definition allows for one month breaks in earnings we cannot guarantee the availability of an interior month without using a two month adjacency period. "Short spells", where no interior month is available for imputation, still exist and the three possible forms these can take are shown as jobs D-F in figure 4. Because short spells consist exclusively of job starts and ends, they cannot be treated in the same way as other starts and ends, and we make a separate adjustment to the FTE for these job-months (discussed at the end of this subsection).

Figure 5 shows the distribution of spell lengths for all spells starting between June 1999 and January 2011, with spell durations over three years pooled. As the figure shows, over 40 percent of starting jobs are short spells, predominantly of one month duration. In addition, a substantial proportion (15-22 percent) of starting job spells have at least one one-month break in them and the prevalence of such gaps rises with observed duration (excluding the over three year group), suggesting that the repair of gaps has an important impact on tenure statistics.[30] Short spells are seasonal, disproportionately starting in December – 12.8 percent of observations (a 53 percent deviation from random) – due to Christmas trading.[31]

[30]See Timmins (2008) for earlier analysis of this phenomenon.

[31]March and November are also overrepresented in short spell start months, capturing 9.9

While short spells represent a large proportion of job starts, they represent less than a third of combined job start and end months (and only 4 percent of all job months). For non-short spell starts and ends, we use interior month earnings divided by interior month FTE to calculate a wage rate. That rate is then used to derive an FTE for the start/end month, which is compared to the FTE calculated from prior steps.[32] If the implied interior month-based FTE is lower, then this becomes the new FTE.

This adjustment tends to systematically overestimate the FTE at the end of job spells, relative to the start, because workers often receive extra payments in their final month of work relating to, eg, accrued leave or wages being paid in arrears. Fortunately, for a significant proportion of job ends, there is an end date filed by the employer with IR (see IR348, appendix B). In these cases, we further adjust the FTE down to the proportion of days worked in the final month of the year where this yields a lower FTE than that previously calculated.[33]

Ends that have not been adjusted through this process, together with all unadjusted short spell months, are then subject to a cap of 0.5 FTE to approximate the average expected period worked during these months. While this adjustment is somewhat arbitrary, it brings the distribution of implied end dates closer to that observed from actual end dates reported to IR.

4.4 Relative incidence of FTE adjustment mechanisms

Table 4 shows the relative frequency of each adjustment technique.[34] The top row of data relates to all job-months pooled and, reading from left to

and 9.1 percent of observations respectively. March seasonality in LEED data is linked to seasonal labour demand in horticulture (Timmins 2009).

[32] In the case where the minimum wage has determined the start month wage, and the minimum wage changes in the next two months, a special adjustment is necessary to ensure that the appropriate minimum wage is still applied in the start month.

[33] The FTE is actually calculated as the proportion of the month multiplied by the minimum wage-adjusted total FTE, L_{it}^{tot}, which results in L sometimes being less than 1/31 (as shown in the following subsection).

[34] This table uses data spanning 180 months from the December 2014 IDI instance of derived tables, where the method included an additional FTE adjustment based on non-wage income share (discussed further in appendix A.3). Affected job-months, accounting for 2.5 percent of total observations (and between 0.6 and 4.4 percent of reported subgroup observations), have been dropped. Because these observations may have been subject to a different (lesser) downward adjustment in the absence of the non-wage component, the data exclusion will only approximate the statistics generated by the method described in the main text.

right, reports statistics on: the share of job-months captured (equal to one for all jobs), the total number of job-months; the proportion of these job-months that have their final FTE determined by each adjustment type; and the resulting average FTE for adjustments of each type. Finally, in the right-most column, the overall average FTE is reported. Below the first row are three mutually exclusive subgroups of job-months which, in turn, are further disaggregated by whether the observation is a start, end or mid-spell month of the job.

These groupings have been chosen to reflect the logic of the adjustment process, isolating observations that are exposed to different permutations of adjustment mechanisms. For example, only multiple job workers are subject to the multi-job adjustment, so excluding single job workers makes it clearer what the effect of the multi-adjustment is, conditional on "eligibility". Further, the multi-job adjustment interacts with a variable number of other technologies depending on whether the month is at the start or end of a job. For clarity, short spells are analysed separately, since they present the unique case where a job-month can be both a spell start and end, resulting in some apparent start months appearing to have "end-month only" technologies applied to them.

Of the approximately 314 million job-month observations, 37.5 percent have a final job-level FTE of less than one, so that the average measured L across all job-months, including unadjusted ones, is 0.791 FTE (top row, table 4). The minimum wage adjustment is the dominant technology that applies in the final data (23.8 percent of observations), followed by the multi-job allocation (7.3 percent), the interior wage adjustment to start and ends, supplied IR end dates and, finally, the capping of previously unadjusted short spells and spell ends.

The minimum wage adjustment dominates because 80 percent of job-months are interior spell months of single job holders (second row of table 4), which are exposed to only this adjustment, coupled with the fact that over a fifth of such job-months are below the expected total earnings for a 40 hour per week minimum wage worker, on average earning just over half (51.6 percent) that figure.

Interior month multiple job holder months are the next most common job-month type (row 5, 8.1 percent of observations). The minimum wage adjustment is important to this group also, affecting 31.1 percent of observations. Minimum wage and multi-job adjustments are not mutually exclusive because the former applies at the worker level and the latter doesn't change the total estimated labour supplied by the worker – instead it allocates that

total across jobs.[35] For simplicity, observations subject to both adjustments have been allocated to the MW column in the table to create useful mutually exclusive categories that apply to all job types. This allocation is apparent from row 5, since all interior month FTEs for multi-job holders *are* affected by the multi-job allocation. For multiple job holders, therefore, the MW column shows the proportion of observations subject to both the minimum wage adjustment and the multi-job allocation (eg, 20.2 percent of multi-job non-short job starts, row 7), and the proportion subject to the multi-job allocation is the sum of the MW and Multi columns (eg, 59.1 (= 20.2 + 38.9) percent for the same subgroup). Consistent with this, the average FTE for job-months subject to both technologies is lower than that where only the multi-job allocation has been applied since the total FTE to allocate between jobs is higher (equal to one) in the latter case.

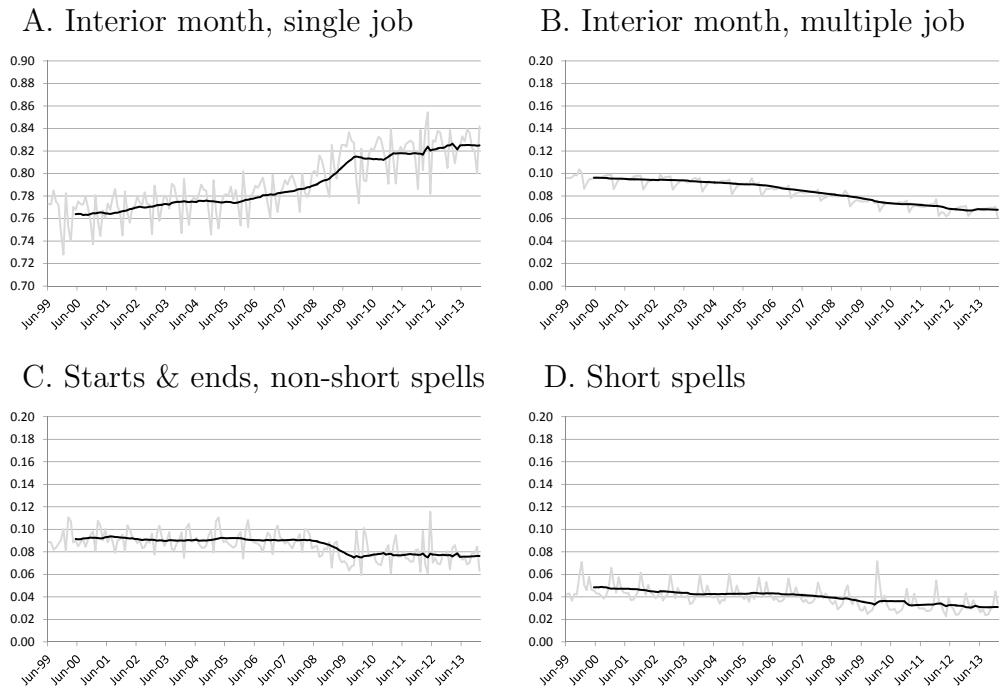
The interior month wage adjustment is an important technology for non-short job starts (41-47 percent of observations, rows 4 & 7) and, to a lesser extent, job ends (24 percent of observations), with the difference consistent with exit payments biasing the calculation for exits. For exits, therefore, actual reported end dates are a more effective adjustment mechanism for both single and multiple job holders (rows 3 & 6). Finally, because there are missing end dates, the FTE cap applies to 21 (11) percent of job ends for single (multiple) job holders.

The FTE cap is the fall-back technology for all short spells (rows 8-10) but, in practice, is used more frequently than non-short spells only in the case of short spell starts (33 percent of those observations). This outcome is primarily due to the relatively low earnings associated with short-spell jobs, so that the minimum wage adjustment applies to between 40 and 51 percent of short spells, coinciding with earnings, on average, below 20 percent of expected earnings at the applicable minimum wage rate.

Figure 6 shows how the composition of job types has changed over time, with a relative increase in single job interior months, from 76 to 82 percent of observations (panel A). At the same time, interior months for multiple job holders, and short spells have both steadily declined as a proportion of total observations – by 2.8 and 1.7 percentage points respectively (panels B & D). The observed sharp downward shift in the relative number of job starts and ends (1.5 percentage points, panel C) is consistent with previous analysis of the impact of the Global Financial Crisis on reducing both job and worker turnover rates in New Zealand (Maré and Fabling 2013).

[35] For single job holders there is no overlap between the technologies – that is, only one is binding.

Figure 6: Composition of job months by job type



The grey line shows actual job-month shares; the solid line is a twelve-month moving average of those shares. Panel A has a different scale from panels B-D, but increments on the y-axis are equivalent across all four graphs, representing two percent of total job-months. The first and last two months of data have been dropped to enable consistent identification of spell starts and ends.

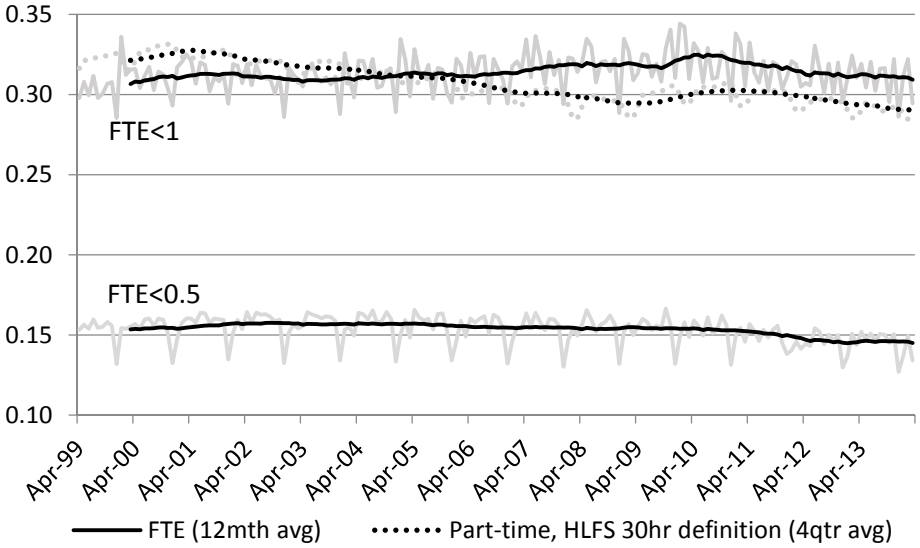
Figure 7 shows the distribution of FTE at the job-month level by adjustment technique – that is, for cases where an adjustment has been made. Across all four methods, there is a tendency for the FTE to be less than one half. In the case of the multi-job only adjustment (panel B), this tendency is balanced by the fact that most workers in this group have only two jobs, introducing a level of symmetry to the FTE distribution – that is, most multiple job workers have a main (FTE$>$ 0.5) and a secondary (FTE$<$ 0.5) job. Aside from IR-notified end dates, there is also a tendency for very low FTE (at or below 0.02) to be relatively uncommon. If IR end dates relied solely on the proportion of days worked, then the minimum observable FTE would be 0.032 (= 1/31) for this category, and we might expect the derived FTE to be more evenly distributed. However, as noted earlier (footnote 33), we pro-rate the minimum-wage adjusted total L, which makes smaller values feasible. End dates are also skewed away from a uniform distribution because of selection – as are other adjustments – since the notified end date is used only where it implies a lower FTE estimate than other applicable adjustment mechanisms.

Figure 8 shows the proportion of part-timers calculated at the worker level over time. The worker level yields a lower estimate of the proportion of

Figure 7: Distribution of FTE by adjustment method

Figures show the proportion of job-months within each adjustment type by 0.02 FTE intervals. By construction, FTE cannot be zero or one for these job-months.

Figure 8: Proportion of part-timers at worker level by month, with HLFS for comparison



Solid grey lines show estimated worker-month shares for workers below half an FTE, and below one FTE; black lines are twelve-month moving averages of those shares. The HLFS proportion of part-time employees (30 hour definition) is shown as a dotted grey line for comparison (black dotted line is four quarter moving average).

observations receiving a labour adjustment than table 4 because the multi-job only adjustment, which affects job-level L, allocates a total FTE of one (at the worker level) across jobs. At the worker level, therefore, the average proportion of adjusted observations declines to 31 percent.[36] Slightly under half of these worker-months relate to individuals whose total FTE is less than one half, again consistent with the aggregation of multiple job holders, who have "secondary" jobs and who may also have had other adjustments applied to their job-level FTE.

Figure 8 also shows the proportion of part-time (less than 30 hour) workers according to the HLFS. The official statistical measure shows a similar proportion of part-time workers as the FTE<1 series, which should not be the case since the HLFS statistics is more akin to FTE<0.75 (assuming an average 40 hour week). This simple comparison suggests that, despite the FTE technology making a credible first-order adjustment to measured labour input, it still overestimates actual L for many workers.

The proportion of interior month, single jobs rises during the Global Financial Crisis (GFC)-induced recession (figure 6, panel A), which would (all else equal) tend to increase the proportion of full-time jobs, since such job-months are the least likely to be adjusted (table 4). However, despite this composition shift, figure 8 clearly shows an upward movement in the proportion of less than full-timers (FTE<1) post-GFC, consistent with adjustment of labour input at the intensive margin for some jobs. This increase in the proportion of part-timers is also apparent in the HLFS statistics.

5 Examples comparing FTE and headcount

We use three simple examples to demonstrate the effect of using the FTE measure compared to a headcount method (ie, treating all workers as $L = 1$). Each example uses a different unit of observation. Firstly, at the job level, we calculate monthly earnings rates demonstrating the effect of FTE-adjustment on the distribution of "wages," particularly for low earning jobs affected by the minimum wage adjustment. Secondly, at the worker level we compare estimated employee labour input shares, by age and sex, showing how differences in the average FTE for women and men systematically affects these

[36] This is approximately equal to one minus the sum of the unadjusted and multi-job only columns in table 4, except that table 4 counts job-months rather than worker-months and, therefore, gives more weight to multi-job holders than the equivalent worker-level statistic.

shares. Finally, at the firm level we compare labour productivity distributions, separately examining the measurement effect of the choice of L on firms that are transitioning into (or out of) employing. This distinction may be material since jobs in firms undergoing employment transition must all start (or end) during the year.

5.1 Earnings rate (wage)

Figures 9-12 show distributions of job-month level earnings rates, defined either as the log of gross earnings (panels A and B, labelled the "raw" earnings rate) or as the log of (gross earnings/FTE) (panel C, the "FTE-adjusted" earnings rate). Each figure shows these distributions for a single year – April 2012 to March 2013 – so that only a single minimum wage rate applies during the period.[37] Mutually exclusive sub-populations are shown in each figure: mid-spell job months (figure 9, 95.1 percent of total FTE); non-short job starts (figure 10, 2.4 percent of total FTE); non-short job ends (figure 11, 1.6 percent of total FTE); and short spell months (figure 12, 0.9 percent of total FTE).[38]

Top panels in each figure show unweighted distributions, while panels B and C are weighted by FTE. Thus, panels A and C (within a figure) contrast the observed earnings rate distribution using a headcount measure with that using the FTE-adjusted measure, where total labour input in each case is calculated in the same manner as the earnings rate (since the raw rate is equivalent to treating all workers as $L = 1$). The inclusion of panel B allows us to examine the effect of two intermediate steps: holding the earnings rate measure constant while varying the weighting (comparing panels A and B); and holding the weighting constant while varying the earnings rate measure (comparing panels B and C). By construction, the FTE-adjusted earnings rate is at least as high as the raw rate, meaning that the distribution always shifts to the right going from panel B to panel C.

For mid-spell jobs (figure 9), panel A shows that over a quarter of monthly earnings are below the level of a full-time minimum wage worker, which (in logs) is calculated as 7.76 for 2013.[39] Since all job-months below this level have an FTE$<$ 1, FTE-weighting the raw earnings rate has the effect of reducing the relative magnitude of the low earnings tail (panel B).

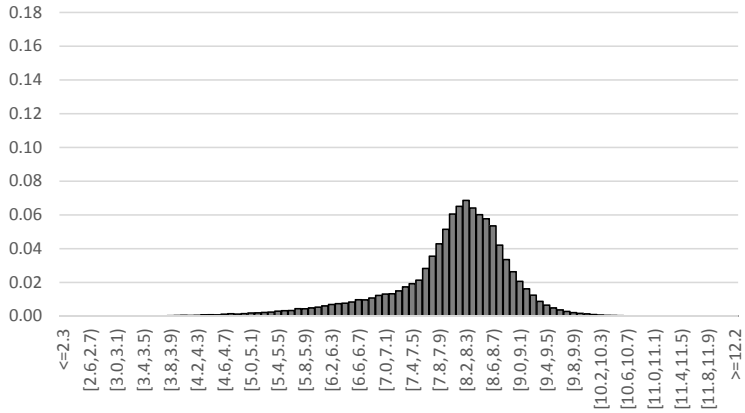
[37] Nominal earnings are used since inflation was low during the year.

[38] These groupings capture 89.1, 4.1, 3.7, and 3.1 percent of total job-months respectively.

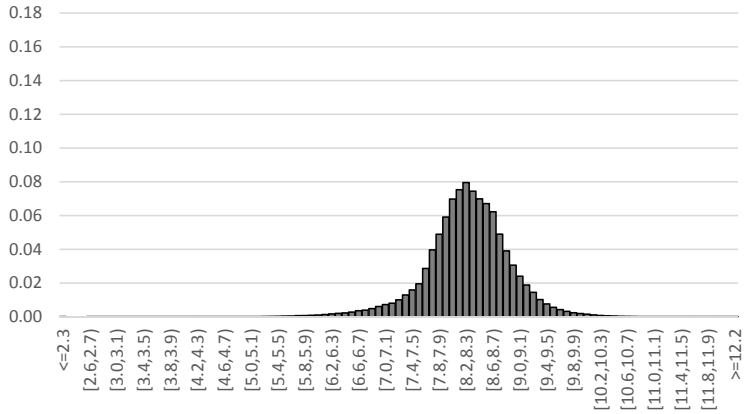
[39] The applicable minimum wage rate is \$13.50/hr so that, using the hours assumption in equation (1), the earnings rate is $\ln(13.5 \times 4.35 \times 40) = 7.76$.

Figure 9: Log earnings rate for mid-spell job-months in 2013

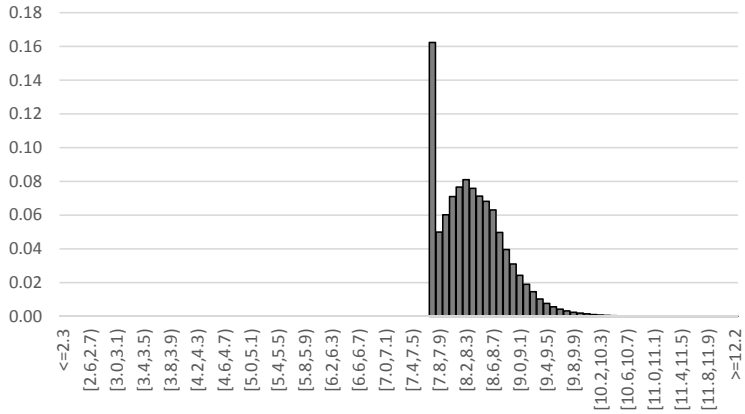
A. Raw earnings (unweighted)



B. Raw earnings (FTE-weighted)



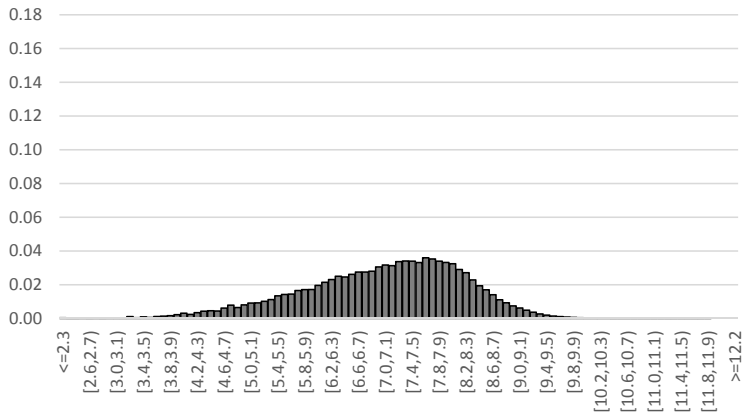
C. FTE-adjusted (FTE-weighted)



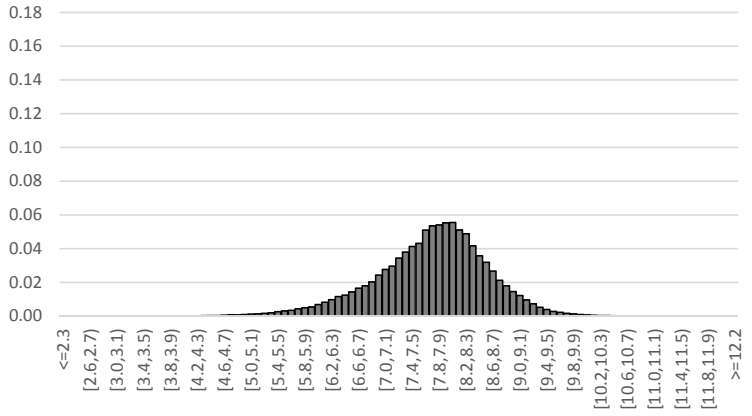
Each panel shows the distribution of the natural log of the earnings rate for mid-spell (non-start/non-end) job-months in the period April 2012 to March 2013. The derivation of the earnings rate differs across the panels. In panels A and B, the earnings rate is the log of gross earnings (equivalent to assuming $L = 1$ at the job level for all workers). Panel A is unweighted and panel B is weighted by FTE. In panel C, the earnings rate is the log of (gross earnings/FTE), and weighted by FTE. In panels A and B earnings below \$10 (above \$200K) are pooled in the bottom (top) grouping. Panel C has analogous groupings on an FTE-adjusted basis. The spike in panel C is due to the minimum wage rate being \$13.50/hr during the year, implying a log of (gross earnings/FTE)= 7.76 under the hours assumption in equation (1).

Figure 10: Log earnings rate for non-short spell starts in 2013

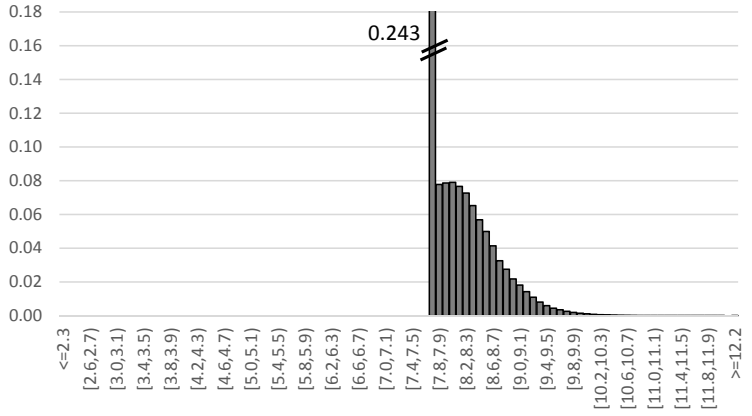
A. Raw earnings (unweighted)



B. Raw earnings (FTE-weighted)



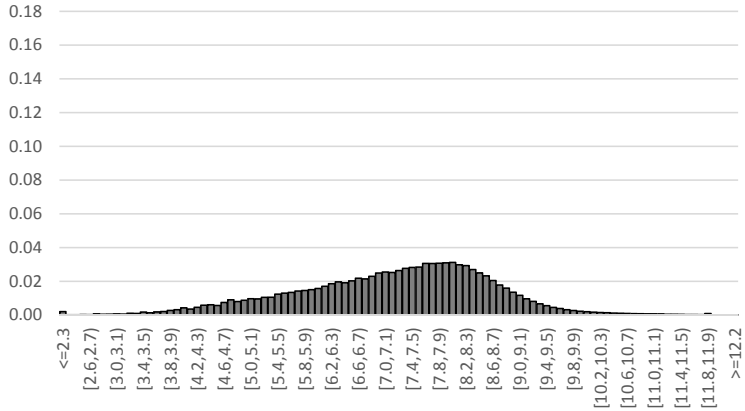
C. FTE-adjusted (FTE-weighted)



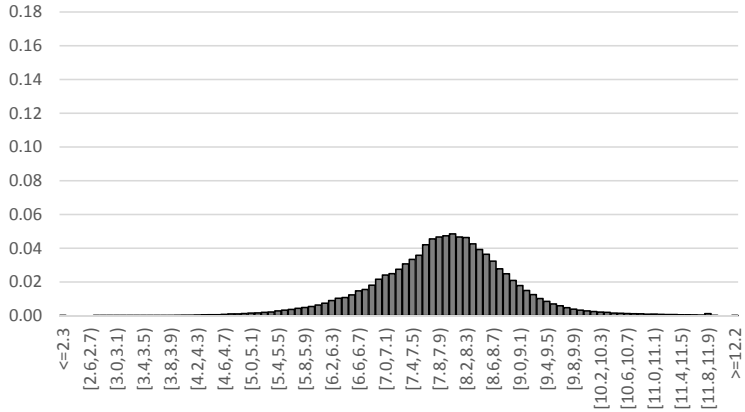
Each panel shows the distribution of the natural log of the earnings rate for non-short spell start months in the period April 2012 to March 2013. See the table note to figure 9 for additional details. The minimum wage-related spike in panel C is truncated to enable the shape of the remainder of the distribution to be observed.

Figure 11: Log earnings rate for non-short spell ends in 2013

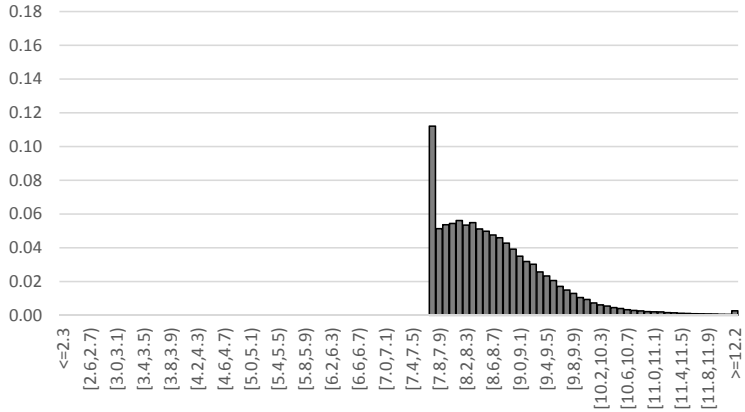
A. Raw earnings (unweighted)



B. Raw earnings (FTE-weighted)



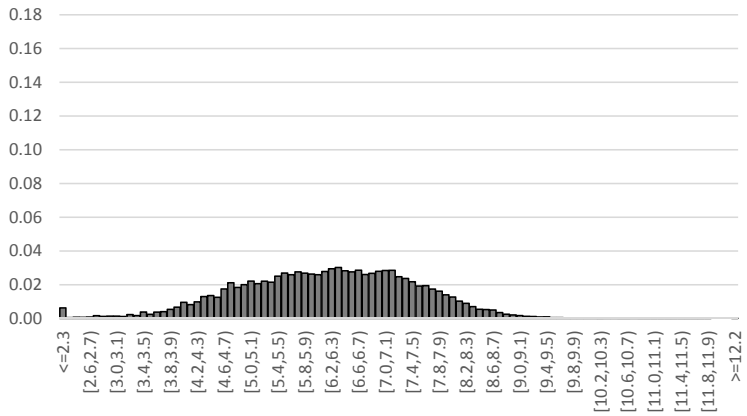
C. FTE-adjusted (FTE-weighted)



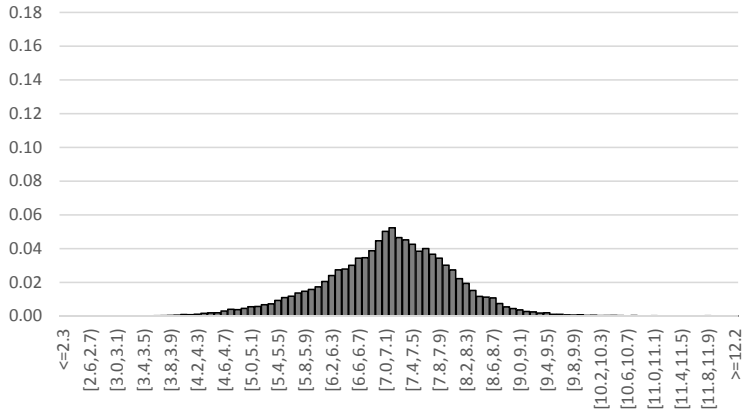
Each panel shows the distribution of the natural log of the earnings rate for non-short spell end months in the period April 2012 to March 2013. See the table note to figure 9 for additional details.

Figure 12: Log earnings rate for short spell job-months in 2013

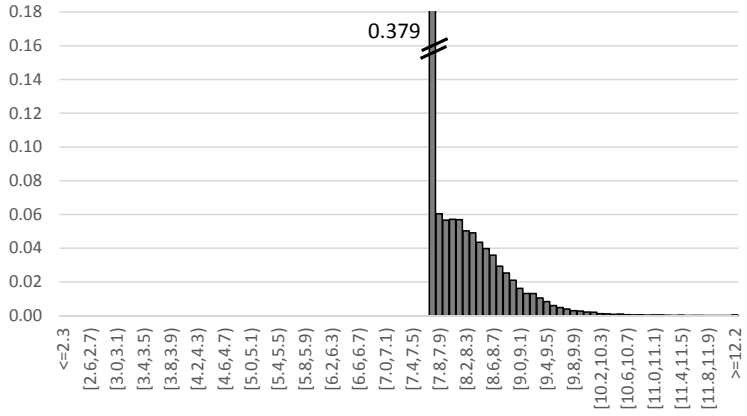
A. Raw earnings (unweighted)



B. Raw earnings (FTE-weighted)



C. FTE-adjusted (FTE-weighted)



Each panel shows the distribution of the natural log of the earnings rate for short spell job-months in the period April 2012 to March 2013. See the table note to figure 9 for additional details. The minimum wage-related spike in panel C is truncated to enable the shape of the remainder of the distribution to be observed.

Measuring the log earnings rate on an FTE-adjusted basis then accumulates the FTE associated with those job months at a single point, as represented by the spike in panel C, since the earnings rate is no longer allowed to be below the minimum wage rate.

For mid-spell months, the distribution of earnings rates to the right of the minimum wage rate is largely unaffected by the choice between headcount and FTE, because the only other technology used is the multi-job adjustment and, on average, only 8 percent of non-MW mid-spell job months are associated with multiple job holders.[40] In summary, the main effect for continuing (mid-spell) jobs is to impose the minimum wage rate.

Low earnings are more prevalent for other types of job month, where between 60 and 88 percent of raw earnings rate observations sit below the floor set by the minimum wage (panel A, figures 10-12). For these subpopulations, the use of additional adjustment mechanisms means that the distribution of the earnings rate shifts noticeably rightward at rates above the minimum wage rate.

Finally, comparing the FTE-adjusted earnings rate distributions across job types, job ends appear to have more density at high values, consistent with exit payments inflating earnings in the last month of jobs. This effect is apparent in both the headcount and FTE-adjusted measures, though is more pronounced in the latter.

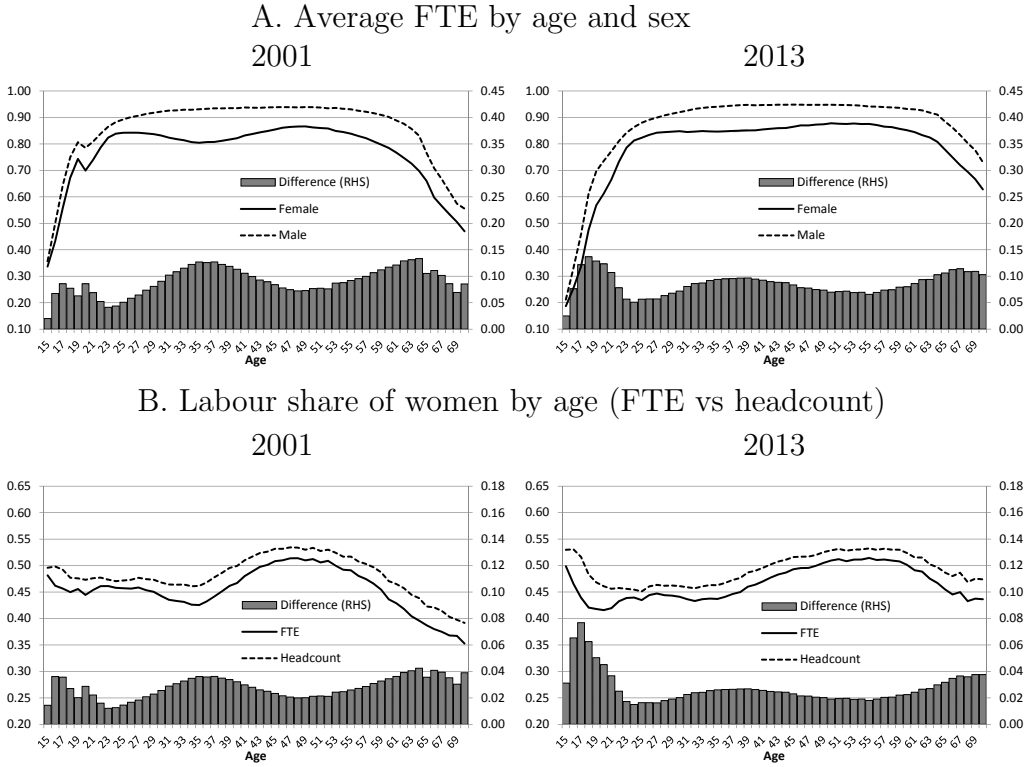
5.2 Relative labour input of female employees

Figure 13 shows the effect, at the worker level, of using FTE to measure L where the research goal is to compare labour input by sex and/or age.[41] Panel A shows the average FTE of female workers compared to male workers by age. At all ages (and across each year) the estimated FTE of women (solid line) is lower than that of men (dashed line). The shaded area shows the percentage point difference between the two lines (using the right-hand scale). Also apparent in 2001, is a dip in average FTE for female in their late 20s through to their early 40s, presumably at least partly related to

[40] Using table 4 (rows 2 & 5), this proportion is calculated as $(0.081 \times 0.689)/(0.795 \times 0.781 + 0.081 \times 0.689)$.

[41] Maré and Hyslop (2006) compare a simpler version of our FTE measure to HLFS statistics by sex, concluding that the tax and survey data sources produce similar aggregate results. In this paper, we take this finding as sufficient evidence that the FTE-based measure is closer to the truth than the headcount measure, and aim to quantify the difference between the two measures.

Figure 13: Effect of FTE on estimated female labour share by age



FTE and headcount are calculated at the worker-month level, and are unweighted averages over a March year (ie, April 2000-March 2001 or April 2012-March 2013). The derived labour share is total L (either FTE or headcount) of x-aged women in a year divided by the total L of all x-aged workers (using the same measure of L). Differences are calculated as the (percentage point) gap between the two labour share lines. Workers younger than 15 and older than 70 are excluded.

childcare. By 2013 this dip is no longer as apparent, though the difference between men and women still shows the corresponding bump with a reduced magnitude compared to 2001.

Also, as expected, the estimated FTE is lower for younger workers of both sexes, despite the lower minimum wage rate for youth in 2001,[42] and for workers past the retirement age of 65. The average FTE for retirement-aged individuals has been rising over time.[43]

Panel B of figure 13 shows how these differences in average FTE translate into female labour shares, calculated as total L for women at a given age

[42] A lower minimum wage rate will yield a higher average estimate of FTE, all else held constant. This is apparent in 2001, where both average FTE lines have a kink, due to the separate youth minimum wage rate applying during the period (and through to 2008). The apparent dip in derived FTE occurs at age 20, which is where the adult rate kicked in (for eleven of the twelve months in the year, see table 3). In 2001, the adult rate was 1.66 times the youth rate. If 19 and 20 year olds have similar earnings ability, the substantial discontinuity in minimum wage rates generates a higher estimated FTE for 19 year olds, compared to 20 year olds.

[43] Comparison of changes in youth FTE over time is difficult given the importance of the removal of the youth-specific minimum wage rate to the calculation.

divided by total L for the all workers of that age.[44] The headcount-based share, therefore, shows relative participation whereas the FTE approach additionally (partially) accounts for relative hours worked. For example, a value of 0.5 would imply the same number of male and female workers under the headcount measure, and the same total labour input under the FTE measure.

The FTE gap for females in their 30s is still apparent in both years (evidenced by the percentage point difference bump). What is also apparent from the labour shares is the relatively lower participation for women also over this age range, particularly in 2001 (ie, the dip in the dashed headcount-based line). Panel B also shows increased relative participation, from 2001 to 2013, of women leading up to and after the retirement age of 65. In percentage point difference terms, the largest gap between the two labour share measures arises for youth under 21 years of age in 2013.

5.3 Labour productivity

Figure 14 shows the effect of the measurement method for employee labour input on estimated labour productivity for a subset of firms.[45] Each panel shows the distribution of firms by productivity level. As with the worker-level example, FTE-based estimates are shown as solid lines; headcount-based estimates are represented by dashed lines; and shaded areas show the percentage point difference in density between the two approaches (using the right-hand scale). Continuers (panel A, and representing 78 percent of observations) are firms that have employees in both the prior and following year; entrants (panel B, 14 percent of observations) have no employees in the prior year; and exiters (panel C, 8 percent of observations) have no employees in the following year.[46]

All three groups show dispersion in labour productivity, which is due to many factors including differences in capital intensity (both tangible and intangible), unmeasured heterogeneity in labour (eg, skills composition), and

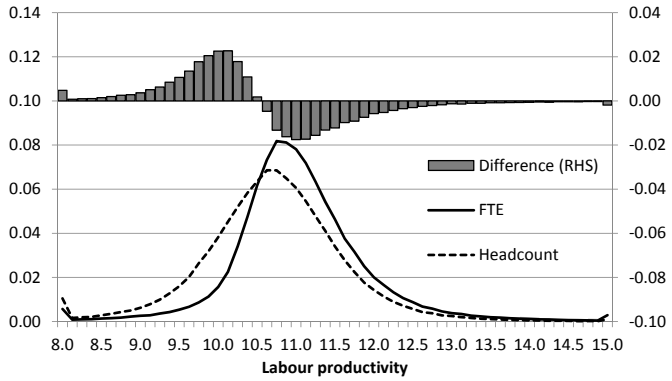
[44] This analysis excludes the self-employed and, therefore, should not be compared to measures of total labour input by sex.

[45] To focus on the effect of employee labour input, we restrict the sample to firms without working proprietors in the year (ie, employee-only firms). Value-added data come from the Fabling-Maré productivity dataset (Fabling and Maré 2015) and is all years (2001-2012) pooled.

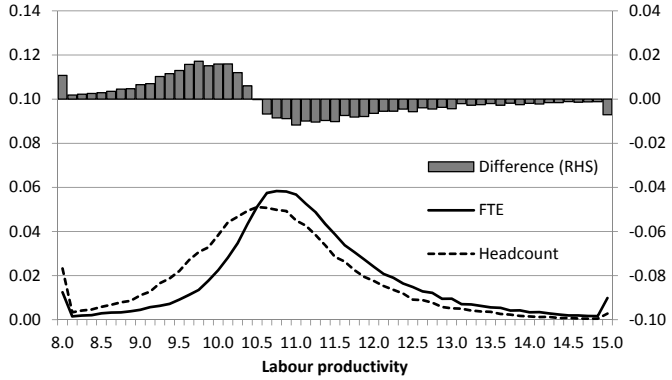
[46] We exclude firms that employee in neither the prior nor the following year so that firms can be categorised cleanly into the three groups. Because labour productivity is measured in logs, negative value-added firms, which are disporportionately exiters, are lost from the sample.

Figure 14: Firm labour productivity by employment transition

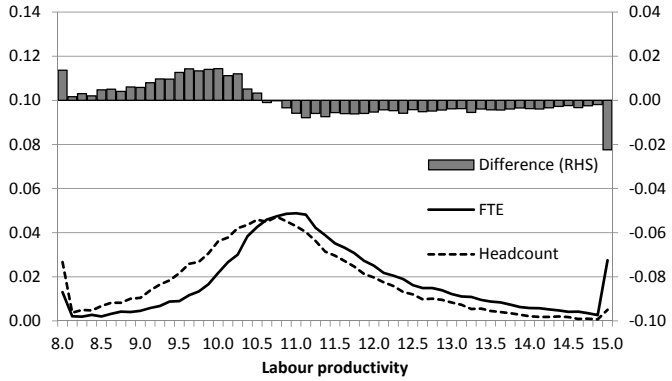
A. Continuers



B. Entrants



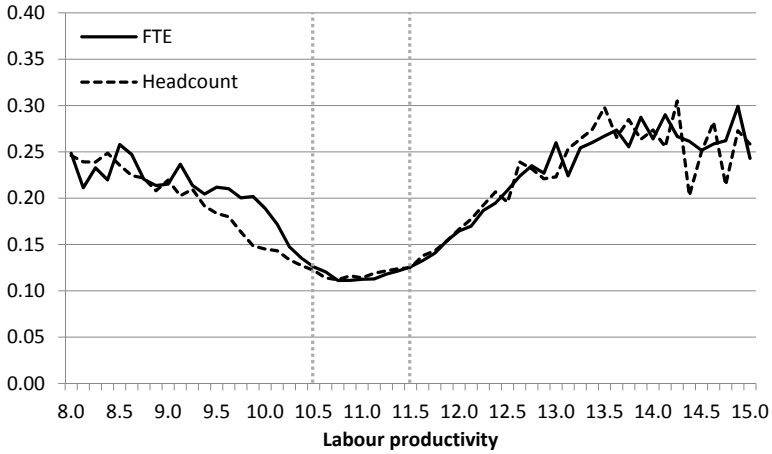
C. Exiters



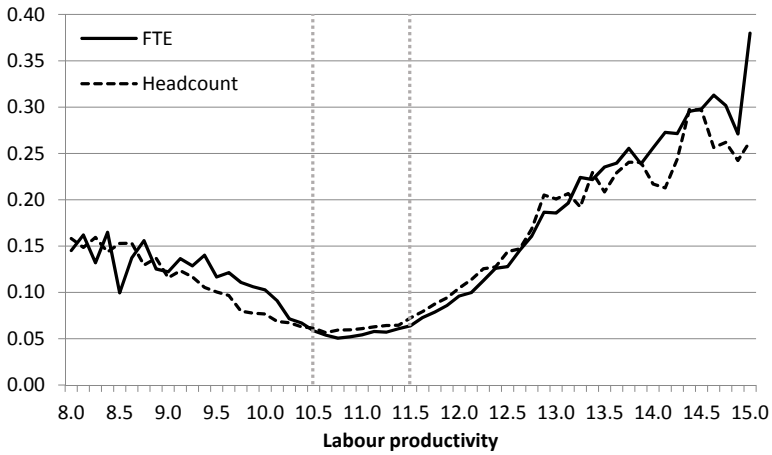
Labour productivity (LP) calculated as ln(value-added/L) where L is either a twelve-month average of employee headcount or FTE. Frequency is calculated at intervals of 0.125, and is pooled in the lowest (highest) category for LP$<$ 8 ($>$ 15). Firms with working proprietors are excluded, as are firms that employ in neither the previous nor the next year. Transition groups are based on employment status: continuers are firms who have employees in both the previous and next year; entrants have no employees in the prior year; and exiters have no employees in the following year.

Figure 15: Share of firms that are entrants/exiters by labour productivity

A. Entrants



B. Exiters



Labour productivity calculated as ln(value-added/L) where L is either a twelve-month average of employee headcount or FTE. Frequency is calculated at intervals of 0.125, and is pooled in the lowest (highest) category for LP<8 (>15). Vertical dotted lines show the interquartile range (ie, 25th-75th) of the FTE-based labour productivity distribution. Firms with working proprietors are excluded, as are firms that employ in neither the previous nor the next year. Transition groups are based on employment status: entrants have no employees in the prior year; and exiters have no employees in the following year.

measurement error. The use of the FTE-based measure reduces this apparent dispersion, at least for continuers and entrants. The largest impact on the distribution (as measured by percentage point differences), is for continuers, which may initially seem counterintuitive. To see why this may be, recall that a large proportion of all jobs are adjusted (37.5 percent) and that the majority of these adjustments are due to the minimum wage and multi-job technologies, which are applicable to mid-spell jobs. Over and above this, continuing firms experience substantial worker turnover, meaning that other (spell start/end) adjustments are also relevant to the measurement of L for these firms. Finally, conditional on adjustment, the average job month FTE is 0.443,[47] which is markedly different from the headcount-based measure.

Since the compression of the labour productivity distribution is stronger for continuing firms, and because these firms make up most of the sample, the apparent *relative* labour productivity of entrants and exiters shifts when the measure of L changes. This effect is shown in figure 15, which shows the share of firms that are entrants (panel A) or exiters (panel B) at each point of the productivity distribution. Vertical dotted lines show the interquartile range (ie, the 25th and 75th percentiles) of the FTE-based labour productivity distribution to give a sense of where the majority of observations lie.[48] The fact that the solid line lies above the dashed line to the left of the 25th percentile of the labour productivity distribution indicates that the FTE-based measure *increases* the relative density of entrants and exiters at relatively low productivity levels.[49]

Finally, the other key point from this graph is also reflected in figure 14. The use of the FTE-based measure substantially increases both the relative and absolute density of exiters with measured labour productivity above a value of 15 (which are pooled in the top category of each figure). In some sense, the headcount measure fares no better, as this method causes a large accumulation of exiters at very low productivity levels (figure 14, panel C), though inferior relative productivity would at least be more theoretically consistent with subsequent exit. These issues probably partly reflect problems measuring value-added for exiting firms.

[47] This can be backed out from table 4, row 1, applying FTE= 1 to the unadjusted workers $(0.443 = (0.791 - 0.625)/(1 - 0.625)$.

[48] Given that the sample is dominated by continuers, the overall distribution is largely inferable directly from figure 14 for both the FTE- and headcount-based measures.

[49] This statement is also true using the headcount-based 25th percentile, as this also lies above 10.

6 Conclusion

In response to the absence of hours information in New Zealand's linked employer-employee data, we have developed a method of deriving an approximation to individual monthly labour input using a plausible set of assumptions about workers, namely that: the statutory minimum wage is observed; multiple-job workers have the same total labour supply as single-job workers; and that hourly wage rates are likely to be constant over adjacent months (at the start and end of jobs). As part of this derivation, we also identify job spells correcting for short (one month) breaks in earnings, which may be useful to researchers analysing job transitions (among other things).

While the resulting adjusted labour input is an overestimate of actual L for many workers, we believe the approximation is substantially superior to a simple headcount approach. Three simple examples comparing the FTE and headcount approach suggest that tangible differences in findings are apparent for economically interesting topics: measurement of wage rates, comparison of workers by gender and age, and comparison of firms by labour productivity. Even without adopting the FTE measure, the associated method identifies – through a series of indicator variables – a subset of job-months that researchers may wish to exclude from any employment analysis, at least as a robustness check on their main results.

Simple tables and metadata (appendix A) have been developed alongside the method, which should enable researchers to use the technology easily. These data are available to any researcher meeting Statistics New Zealand's criteria for access. To further improve usability, the tables are also fully integrated with other technologies that have been developed by the authors: notably the permanent enterprise number (Fabling 2011); estimates from the (log wage) two-way fixed effects model in Maré et al. (2015), commonly used to identify a proxy measure for worker skill; and the productivity dataset on the LBD (Fabling and Maré 2015).

References

Fabling, R. (2009). A rough guide to New Zealand's Longitudinal Business Database. Global COE Hi-Stat Discussion Papers No. 103, Institute of Economic Research, Hitotsubashi University.

Fabling, R. (2011). Keeping it together: Tracking firms in New Zealand's Longitudinal Business Database. Working Paper 11-01, Motu Economic and Public Policy Research.

Fabling, R. and D. C. Maré (2015). Production function estimation using New Zealand's Longitudinal Business Database. Working Paper 15-15, Motu Economic and Public Policy Research.

Fabling, R. and L. Sanderson (2014). Productivity distributions in New Zealand: The dangers of international comparisons. mimeo.

Fabling, R. and L. Sanderson (2015). A rough guide to New Zealand's Longitudinal Business Database (2nd edition). Working Paper forthcoming, New Zealand Treasury.

Maré, D. C. and R. Fabling (2013). The incidence and persistence of cyclical job loss in New Zealand. IZA Discussion Papers 7745, Institute for the Study of Labor.

Maré, D. C. and D. Hyslop (2006). Worker-firm heterogeneity and matching: An analysis using worker and firm fixed effects estimated from LEED. LEED research reports, Statistics New Zealand.

Maré, D. C., D. Hyslop, and R. Fabling (2015). Firm productivity growth and skill. Working Paper forthcoming, Motu Economic and Public Policy Research.

Maré, D. C., L. Sanderson, and R. Fabling (2014). Earnings and employment in foreign-owned firms. Working Paper 14-10, Motu Economic and Public Policy Research.

Statistics New Zealand (2015). Labour market statistics: March 2015 quarter. Statistics New Zealand: Wellington.

Timmins, J. (2008). Why are there so many short jobs in LEED? An analysis of job tenure using LEED. LEED research reports, Statistics New Zealand.

Timmins, J. (2009). Seasonal employment patterns in the horticultural industry. LEED research reports, Statistics New Zealand.

Tables

Table 1: Proportion of working proprietors by number of businesses and employee status elsewhere

WP	Employee elsewhere		
businesses	No	Yes	Total
1	0.663	0.264	0.927
2	0.054	0.013	0.067
3	0.004	0.001	0.005
4+	0.001	0.000	0.001
Total	0.722	0.278	1.000

All years (2000-2013) pooled. Whether WPs is an employee elsewhere is calculated assuming a common 31st March year end.

Table 2: Proportion of working proprietor transition years

	Number of	Proportion that are transitions				Total
	WP-pent obs.	Exit only	Entry only	Both	Total	excl. multi
2000	445,125		N/A			
2001	443,082	0.131	0.102	0.034	0.267	0.206
2002	427,767	0.129	0.102	0.033	0.264	0.207
2003	421,311	0.129	0.111	0.038	0.278	0.220
2004	414,687	0.135	0.111	0.042	0.288	0.230
2005	404,160	0.126	0.111	0.044	0.281	0.227
2006	399,888	0.123	0.117	0.044	0.284	0.232
2007	392,388	0.121	0.109	0.043	0.273	0.226
2008	386,643	0.128	0.104	0.047	0.279	0.233
2009	370,095	0.125	0.096	0.043	0.264	0.222
2010	355,797	0.114	0.094	0.041	0.249	0.212
2011	350,580	0.120	0.098	0.044	0.263	0.225
2012	343,791	0.122	0.100	0.048	0.270	0.230
2013	332,235		N/A			

Transitions are observations where the WP was not observed at the firm in the prior and/or following year. The both category is both an exit and an entry transition and, therefore, an isolated single year where the WP works in the firm. The final column excludes WP with multiple businesses from the transition group, reflecting the proportion of the observations that are affected by the imposed transition cap on L. It is impossible to consistently calculate entry (exit) in the first (last) year, as discussed in the main text, and so transition statistics are excluded.

Table 3: Minimum wage rates

From	To	Minimum wage rate (m, \$\hr) Adult	Youth	Minimum age for adult rate
Apr-99	Feb-00	7.00	4.20	20
Mar-00	Feb-01	7.55	4.55	20
Mar-01	Mar-02	7.70	5.40	18
Apr-02	Mar-03	8.00	6.40	18
Apr-03	Mar-04	8.50	6.80	18
Apr-04	Mar-05	9.00	7.20	18
Apr-05	Mar-06	9.50	7.60	18
Apr-06	Mar-07	10.25	8.20	18
Apr-07	Mar-08	11.25	9.00	18
Apr-08	Mar-09	12.00	–	16
Apr-09	Mar-10	12.50	–	16
Apr-10	Mar-11	12.75	–	16
Apr-11	Mar-12	13.00	–	16
Apr-12	Mar-13	13.50	–	16
Apr-13	Mar-14	13.75	–	16
Apr-14	Mar-15	14.25	–	16

New Zealand currently does not have a minimum wage rate for workers under the age of 16. We assume, therefore, that the adult rate applies to these individuals. At present, there is a "starting-out" wage rate ($11.80) which may apply to youth aged under 20, either in the first six months of work or during periods of training. In addition, there is a "training" minimum wage (also $11.80) for adults in recognised industry training. Both these minimum wage rates have been ignored in the analysis since it is not possible to identify the workers who meet the training criteria.

Table 4: Job-months by adjustment method used, Apr-99 to Mar-14 (180 months) pooled

		Job-month		Proportion of within-group job-months						Average FTE				
Group		Share	Total	Unadj.	MW	Multi	Inter.	IR date	Cap	MW	Multi	Inter.	IR date	Overall
All jobs														
		1.000	313,645,860	0.625	0.238	0.073	0.029	0.019	0.016	0.443	0.462	0.441	0.313	0.791
Non-short spell, single job holders														
Start	End													
N	N	0.795	249,270,318	0.781	0.219	0.000	0.000	0.000	0.000	0.516	N/A	N/A	N/A	0.894
N	Y	0.028	8,673,996	0.000	0.182	0.000	0.240	0.364	0.214	0.180	N/A	0.515	0.373	0.399
Y	N	0.028	8,761,014	0.153	0.375	0.000	0.472	0.000	0.000	0.300	N/A	0.525	N/A	0.513
Non-short spell, multiple job holders														
Start	End													
N	N	0.081	25,513,428	0.000	0.311	0.689	0.000	0.000	0.000	0.306	0.529	N/A	N/A	0.460
N	Y	0.014	4,534,896	0.000	0.164	0.217	0.243	0.264	0.112	0.148	0.203	0.224	0.287	0.254
Y	N	0.015	4,592,232	0.000	0.202	0.389	0.409	0.000	0.000	0.225	0.398	0.302	N/A	0.323
Short spells														
Start	End													
N	Y	0.010	3,043,056	0.000	0.403	0.162	0.000	0.238	0.197	0.164	0.144	N/A	0.275	0.253
Y	N	0.010	3,064,248	0.000	0.513	0.156	0.000	0.000	0.331	0.198	0.174	N/A	N/A	0.294
Y	Y	0.020	6,192,669	0.000	0.473	0.246	0.000	0.119	0.162	0.153	0.127	N/A	0.137	0.201

Adjustment technologies are: minimum wage (MW); multiple job holder (Multi); interior wage (Inter.); notified end date (IR date); and capped (Cap). Remaining observations are unadjusted (Unadj.), ie FTE=1. Average FTE is not reported for capped job-months since this is 0.5 by construction. Statistics relate to the December 2014 IDI instance of the derived tables, where the method included an additional FTE adjustment based on non-wage income share (discussed in appendix A.3). Affected job-months, accounting for 2.5 percent of total observations (and between 0.6 and 4.4 percent of the reported subgroups), have been dropped. Because these observations may have been subject to a different (lesser) downward adjustment in the absence of the non-wage component, the data exclusion will only approximate the statistics generated by the method described in the main text.

Appendices

A Data dictionaries for derived tables

A.1 IDI (IDI_Sandpit) tables

The WP and FTE algorithms each generate a single PENT × individual × t table, where time (t) is annual for working proprietors and monthly for jobs. Firm identifiers in all tables are based on permanent enterprise numbers (Fabling 2011). The jobs table includes data derived from estimating the two-way fixed effects model described in Maré, Hyslop, and Fabling (2015). Given the large size of the job month table, we also provide a simple aggregation of jobs to the PENT × month level for convenience.

IDI_Sandpit rules require the name of the researcher compiling the tables to be appended to the table name. This name may vary across updates, as will the IDI instance.

A.2 LBD (IBULDD_Research_Datalab) tables

The LBD contains two aggregations of the IDI tables – one at PENT × year level, and one at PBN (plant) × month level. The first of these tables incorporates estimates from Maré et al. (2015), and is a companion to the productivity dataset of Fabling and Maré (2015). These two tables provide substitutes for LEED variables that were previously available in:

1. `fact_lbf_enterprise_year.leed_rme_as_at_15th_no_WP` replaced by `pent_year_L_IDI_20141205.rme_no_WP` (or `fte` if labour adjustment method preferred)[50]
2. `fact_leed_enterprise_year.WP_count_nbr` replaced by `pent_year_L_IDI_20141205.WP`

[50]The old LEED RME measure imputes missing start and end dates which means that, on average, half of all start months are not counted as part of the measure (since employment is measured "as at" mid-month). This treatment implies that the old RME measure sometimes lies below the new RME measure – and, potentially, also below the FTE measure – if workers join or leave a firm during the year. The old and new counts will also differ because of differences in the definition of working proprietors.

3. `load_lbf_fact_pbn_employee_count.leed_employee_count` replaced by `pent_pbn_month_L_IDI_20141205.employee_count` (or `fte` if labour adjustment method preferred).[51]

At the firm level, non-standard (ie, non-March) balance dates can mean that the EMS data doesn't include all twelve months of a year, at least in the first and last `dim_year_key`.[52] The firm-level tables account for this by averaging the FTE (and RME) count over only the number of months where EMS data exists,[53] which is the method formerly used for LEED variables in the LBD.

While labour variables all derive from the same IDI instance, the labour dataset includes all employing businesses, whereas the productivity dataset includes PENT-years only where usable productivity components are available.[54] The labour dataset should be used by researchers not interested in productivity, or who are interested in assessing the undercoverage of the productivity dataset, either in terms of employing firms or employment.

The productivity dataset WP count may differ from the labour dataset in years where WP transitions are unknown. Since the productivity dataset starts in the 200103 `dim_year_key`, this issue arises only in the last year of data. This is because both productivity components and WP counts rely on annual tax filing, meaning that WP data post-dating the last productivity dataset year may not exist. Working proprietor counts that may differ will have non-zero `WP_unknown_trans` in the labour table. The productivity dataset methodology makes a firm-level adjustment, analogous to the WP-level transition adjustment, based on an estimate of whether this is the last year of business activity.[55]

[51] The old LEED-based employee count in this table included working proprietors receiving PAYE income, whereas the new employee count (and FTE) excludes WPs. Additionally, the old method double-counted employees assigned to multiple PBNs during the month, and the new table counts these individuals only once by assigning them to their minimum PBN in the month.

[52] For example, a firm with a (rare) December balance date will have only nine months of EMS data for the 200003 `dim_year_key`, ie April 1999-December 1999.

[53] Here we mean exists for all firms that have employees, not whether the firm in question has filed in the month. EMS months in the tail of the data, where coverage is partial, are treated as not existing.

[54] There also exist minor differences in the population of workers used for the fixed effects estimates. The labour measure assumes that workers with missing age information are adults, whereas the estimates of Maré et al. (2015) require both age and gender to be known.

[55] This adjustment is not applied in the labour dataset because identification of business exit relies on firm-level tax (GST) and survey data, which is not full coverage. In addition, most IDI users do not have access to business tax data which inhibits its inclusion in the

A.3 Current (December 2014) table deviations from the described method

The current (December 2014) instance of the tables do not implement the transition-based adjustment for WPs. Instead, the productivity dataset halves the WP count in the firm's apparent start and end year (rather than the owners start and end year). This alternative approach results in inconsistent WP counts across the labour and productivity tables in all years. It may also undervalue the input of some WPs.[56]

Two further deviations from the described method also occur in the current (December 2014) instance:

1. IR4S income counts towards WP identification only in the specific years that it exceeds the (real) threshold value, rather than counting in all years provided the individual *ever* receives IR4S income above the cut-off.

2. An additional FTE adjustment based on non-wage income share is included which, in essence, treated non-wage income as an additional "job" when determining the multi-job allocation.[57] FTE that are ultimately determined by this adjustment have the `used_nonwage` flag set to one.

Future versions of the tables will be consistent with the labour measurement method described in the main text. Further, the productivity methodology will also be made consistent with this approach by no longer adjusting WP counts for the first and last year of *business* operations, except where the individual-level transition-based adjustment is not possible (as outlined above).

IDI-located datasets. The methodology described in the main text overcomes these issues by relying exclusively on what is observable at the individual level.

[56] For example, if a WP exits one business and starts another in the same year, the approach used in the current tables is to assume WP input is a half in each business in the labour measurement code, since the individual owns and works in two firms in the year. Subsequently, in the productivity code, WP labour input is halved again because it is the first year of operation in one of the firms and the last year of operation in the other.

[57] Non-wage income was subject to a discount factor of 0.9, relative to wage income. The non-wage adjustment has been dropped because the MW adjustment is generally more likely to determine the FTE of individuals with non-wage income, except in cases where the ratio of non-wage to total earnings appears implausibly high. In such cases, it is difficult to justify the non-wage adjustment.

IDI_Sandpit.clean_read_IR.pent_emp_mth_FTE_IDI_20141205_RFabling

Key	Variable	Format	NULLs	Description
⋆	pent	char(10)	N	Permanent enterprise number (firm id) in format "EN" followed by 8-digit number
⋆	snz_uid	int	N	Individual (worker) id
⋆	dim_month_key	int	N	Month in format YYYYMM
	min_pbn_nbr	char(10)	N	Permanent Business Number (plant id) in format "PB" (or "PX") followed by 8-digit number. Minimum PBN number that worker has been allocated to in month
	max_pbn_nbr	char(10)	N	Maximum PBN number that worker has been allocated to in month
	age	float	Y	Age in years (monthly increments), derived from birth month recorded in personal details table
	sex_code	char(1)	Y	Gender as recorded in personal details table ("F"=female; "M"=male)
	gross_earn	decimal(13,2)	N	Total gross earnings from EMS wage & salary in job-month
	fte	float	N	Derived full-time equivalent (FTE) measure of labour input
	n_jobs	int	N	Number of jobs held during the month
	spell_start	tinyint	N	Binary indicator = 1 if month is the start of a job spell (0 otherwise)
	spell_end	tinyint	N	Binary indicator = 1 if month is the end of a job spell (0 otherwise)
	short_spell	tinyint	N	Binary indicator = 1 if month is part of a short spell, ie a job spell with no interior months (0 otherwise)
	mw_binding	tinyint	N	Binary indicator = 1 if earnings in month are at or below the minimum wage threshold (0 otherwise)
	used_adjacent	tinyint	N	Binary indicator = 1 if adjacent interior month wage used to determine FTE (0 otherwise)
	used_end_date	tinyint	N	Binary indicator = 1 if job end date supplied to IR used to determine FTE (0 otherwise)
	used_halve_unadj	tinyint	N	Binary indicator = 1 if spell start/end & FTE set to 0.5 because no other adjustment mechanism applied (0 otherwise)
	xb	float	Y	Measure of the observable worker characteristic (age and gender) component of log wages estimated from the two-way fixed effects model in Maré et al. (2015)

The current (December 2014) instance of this table, which is used to derive other tables in this collection, applies an additional adjustment based on benefit receipt. As a consequence, the table includes two additional variables – `has_nonwage` and `used_nonwage` – with the latter indicating that the (now-defunct) non-wage adjustment determines the FTE.

IDI_Sandpit.clean_read_IR.pent_mth_FTE_IDI_20141205_RFabling

Key	Variable	Format	NULLs	Description
⋆	pent	char(10)	N	Permanent enterprise number (firm id) in format "EN" followed by 8-digit number
⋆	dim_month_key	int	N	Month in format YYYYMM
	total_gross_earn	decimal(13,2)	N	Total gross earnings from EMS wage & salary paid by firm in month
	total_fte	float	N	Total derived full-time equivalent (FTE) measure of labour input
	n_employees	int	N	Headcount measure of paid employees
	n_spell_start	tinyint	N	Number of job spells starting in month
	n_spell_end	tinyint	N	Number of job spells ending in month

IDI_Sandpit.clean_read_IR.pent_WP_yr_IDI_20141205_RFabling

Key	Variable	Format	NULLs	Description
⋆	pent	char(10)	N	Permanent enterprise number (firm id) in format "EN" followed by 8-digit number
⋆	snz_uid	int	N	Individual (working proprietor, WP) id
⋆	dim_year_key	int	N	Financial (balance date) year in format YYYY03, allocated to "closest" March year
	has_wage_inc	tinyint	N	Binary indicator = 1 if WP was paid EMS wage & salary by firm during month (0 otherwise)
	has_nonwage_inc	tinyint	N	Binary indicator = 1 if WP received non-EMS income from firm during month (0 otherwise)
	multi_WP	tinyint	N	Binary indicator = 1 if individual is a WP of multiple firms during year (0 otherwise)
	WP_prior_year	tinyint	Y	Binary indicator = 1 if individual is a WP of this firm in the prior year (0 otherwise)
	WP_next_year	tinyint	Y	Binary indicator = 1 if individual is a WP of this firm in the next year (0 otherwise)
	adj_WP_count	float	Y	Headcount measure of WP labour input adjusted for multiple business ownership or WP transition in year

The current (December 2014) instance of this table, which is used to derive other tables in this collection, does not apply the transition-based adjustment described in the main text, nor the "ever-threshold" rule for company owners. As a consequence of the former exclusion, three variables – `multi_WP`, `WP_prior_year` and `WP_next_year` – are not present in the table. The variable `multi_WP` can be inferred by observing whether `adj_WP_count` is one.

ibuldd_research_datalab.[STATSNZ\RFabling].pent_pbn_month_L_IDI_20141205

Key	Variable	Format	NULLs	Description
⋆	pent	char(10)	N	Permanent enterprise number (firm id) in format "EN" followed by 8-digit number
⋆	pbn_nbr	char(10)	N	Permanent Business Number (plant id) in format "PB" (or "PX") followed by 8-digit number
⋆	dim_month_key	int	N	Month in format YYYYMM
	dim_year_key	int	N	Financial (balance date) year in format YYYY03, allocated to "closest" March year
	fte	float	N	Total full-time equivalent (FTE) measure of labour input at PBN. Workers "at" multiple PBNs counted only at their minimum PBN number
	employee_count	int	N	Headcount measure of employees at PBN

ibuldd_research_datalab.[STATSNZ\RFabling].pent_year_L_IDI_20141205

Key	Variable	Format	NULLs	Description
⋆	pent	char(10)	N	Permanent enterprise number (firm id) in format "EN" followed by 8-digit number
⋆	dim_year_key	int	N	Financial (balance date) year in format YYYY03, allocated to "closest" March year
	fte	float	N	Average monthly full-time equivalent (FTE) measure of labour input during year
	WP	float	N	Total adjusted working proprietor (WP) count
	WP_unknown_trans	float	N	Total WP count where the WP transition-based adjustment is not implementable
	rme_no_WP	float	N	Average monthly headcount measure of employees during year, known previously as rolling mean employment (RME) excluding WPs
	total_gross_earn	decimal(13,2)	N	Total gross earnings from EMS wage & salary paid by firm in year
	ffe	float	Y	Firm fixed effect estimated from the two-way fixed effects model of log wages in Maré, Hyslop, and Fabling (2015), hereafter the MHF model
	fe_group	int	Y	Identifier of connected groups of firms from the MHF model. That is, firms with the same group identifier are connected, possibly indirectly, by employee movements between firms
	fte_with_wfe	float	N	Average monthly FTE of workers with estimated worker fixed effect from the MHF model (ie, workers with age and gender information)
	avg_wfe	float	Y	FTE-weighted average of the worker fixed effects estimated from the MHF model
	avg_xb	float	Y	FTE-weighted average of the observable worker characteristics (age and gender) component of log wages estimated from the MHF model

The current (December 2014) instance of this table does not make use of transition-based adjustments to calculate the WP count. As a consequence, the variable `WP_unknown_trans` is not present in the table.

B Inland Revenue forms

Inland Revenue
Te Tari Taake

Employer monthly schedule

For help, refer to notes on employer deductions form.

Employer's IRD number **1**

Period ended **2**

This schedule is due

Gross earnings and/or schedular payments	Earnings and/or schedular payments not liable for ACC earners' levy	3 PAYE and/or schedular tax deductions	4 Child support deductions	5 Student loan deductions	6 KiwiSaver deductions	7 Net KiwiSaver employer contributions

Employee name and IRD number

Surname | First name(s) | IRD number | Tax code | Employment start and/or finish date: Start (Day Month Year), Finish (Day Month Year)

Tick if lump sum payment made and taxed at lowest rate | CS code

$ 00	$ 00	$	$	$	$	$

Surname | First name(s) | IRD number | Tax code | Employment start and/or finish date: Start (Day Month Year), Finish (Day Month Year)

Tick if lump sum payment made and taxed at lowest rate | CS code

$ 00	$ 00	$	$	$	$	$

Surname | First name(s) | IRD number | Tax code | Employment start and/or finish date: Start (Day Month Year), Finish (Day Month Year)

Tick if lump sum payment made and taxed at lowest rate | CS code

$ 00	$ 00	$	$	$	$	$

Surname | First name(s) | IRD number | Tax code | Employment start and/or finish date: Start (Day Month Year), Finish (Day Month Year)

Tick if lump sum payment made and taxed at lowest rate | CS code

$ 00	$ 00	$	$	$	$	$

Total gross earnings and/ or schedular payments	**Total earnings not liable for ACC earners' levy**	**3 Total PAYE and/or schedular tax**	**4 Total child support**	**5 Total student loan**	**6 Total KiwiSaver deductions**	**7 Total Net KiwiSaver employer contributions**
$ 00	$ 00	$	$	$	$	$

If your correct daytime phone number is not shown below, print in the box

(   )

If your correct contact person's name is not shown below, print in the box

Declaration

I declare that the information given in this return is true and correct.

Signature

/ /
Date

OFFICE USE ONLY

Operator code | Corresp. indicator | Payment attached | Return cat.

Inland Revenue
Te Tari Taake

Income Tax Act 2007

Company shareholders' details

1 April 2011 to 31 March 2012

IR 4S
2012

- **Read the notes to Questions 29 and 41 on pages 31 and 49 of the IR 4 guide.**
- **Attach this form to the top of page 5 of the company's IR 4 income tax return.**

- Print the details for all shareholders, directors and relatives of shareholders who received remuneration (with no PAYE deducted), or loans from the company, or were attributed a loss by the company.
- For group companies which shared loss offsets and subvention payments, print these amounts in Boxes 41F and 41G.

Company name

IRD number
(8 digit numbers start in the second box 1 2 3 4 5 6 7 8)

- **Complete the following details for each shareholder, director or relative.** See page 49 of the IR 4 guide.
- Please show the shareholder's, director's or relative's IRD number for amounts in Boxes 41B to 41E and the associated company's IRD number for Boxes 41F and 41G on a separate line.

41A IRD number

41B Attributed loss $, , . –

41C Remuneration with no PAYE deducted $, , .

41D Value of loans from the company $, , .

41E Current account balance $, , . (Tick one) CR DR

41F Loss offsets (put a minus sign in the last box to show losses **claimed**) $, , .

41G Subvention payments (put a minus sign in the last box to show subvention payments **made**) $, , .

41A IRD number

41B Attributed loss $, , . –

41C Remuneration with no PAYE deducted $, , .

41D Value of loans from the company $, , .

41E Current account balance $, , . (Tick one) CR DR

41F Loss offsets (put a minus sign in the last box to show losses **claimed**) $, , .

41G Subvention payments (put a minus sign in the last box to show subvention payments **made**) $, , .

41A IRD number

41B Attributed loss $, , . –

41C Remuneration with no PAYE deducted $, , .

41D Value of loans from the company $, , .

41E Current account balance $, , . (Tick one) CR DR

41F Loss offsets (put a minus sign in the last box to show losses **claimed**) $, , .

41G Subvention payments (put a minus sign in the last box to show subvention payments **made**) $, , .

Inland Revenue
Te Tari Taake

Income Tax Act 2007, Tax Administration Act 1994

Income tax return

Partnerships and look-through companies (LTCs)

IR 7 2012

Read the notes on page 4 to help you complete this return.

1 April 2011 to 31 March 2012

If the partnership's or LTC's name and IRD number are shown correctly above, go straight to Question 3.

1. If the partnership's or LTC's IRD number is **not** shown above, print it in Box 1.
(8 digit numbers start in the second box. 1 2 3 4 5 6 7 8)
1

2. If the correct partnership or LTC name is **not** shown above, print it in Box 2.
2

3. If the partnership's or LTC's **trading name** has **changed** or is **not** shown below, print it in Box 3.
3

4. If the correct **postal** address is **not** shown above, and is different to the street address below, print the full address in Box 4. Don't print your tax agent's address here. See notes on page 4.
4
Please put street address or PO Box number **above** and suburb, box lobby or RD and town, city or region **below**

5. If the correct **street** address is **not** shown below, print it in full in Box 5.
5
Please put street address above and suburb or RD and town, city or region **below**

6. Print your business industry classification (BIC) code in Box 6. See notes on page 4.
6

7. If the correct daytime phone number is **not** shown below, print it in Box 7. See notes on page 4.
7

8. Is this the partnership's or LTC's **first** return?
No ○ Go to Question 9.
Yes ○ Print the date the partnership or LTC began in Box 8.
8 Day Month Year

9. Has the partnership or LTC ceased?
No ○ Go to Question 10.
Yes ○ See notes on page 4.

If you have no gross income or claimable losses, and no disclosures to make at Question 26, see the note on page 4 about nil returns. Then complete and sign the declaration on page 3 of this return. You don't need to answer any further questions.

Income

10. Did the partnership or LTC receive any **schedular payments**?
No ○ Go to Question 11. Yes ○ See page 4 of the guide. Print the totals here.
Total withholding tax deducted **10A** $
Total gross schedular payments **10B** $

11. Did the partnership or LTC have any **New Zealand interest** paid or credited to it?
No ○ Go to Question 12. Yes ○ See pages 4 to 6 of the guide. Print the totals here. **Keep any certificates.**
Total RWT **11A** $
Total gross interest—if a loss, put a minus sign in the last box **11B** $

12. Did the partnership or LTC have any **New Zealand dividends** paid or credited to it, or did the partnership or LTC receive shares instead of dividends? Include any dividends from partnerships, LTCs or trusts.
No ○ Go to Question 13. Yes ○ See pages 6 and 7 of the guide. Print the totals here. **Keep any statements.**
Total dividend imputation credits **12** $
Total dividend RWT and payments for foreign dividends **12A** $
Total gross dividends **12B** $

13. Did the partnership or LTC receive any **taxable distributions from a Māori authority**?
No ○ Go to Question 14. Yes ○ See pages 8 and 9 of the guide. Print the totals here. **Keep your Māori authority distribution statements.**
Total Māori authority credits **13A** $
Total Māori authority distribution **13B** $

14. Did the partnership or LTC receive any **income from another partnership**? (Exclude any income/losses received you have included at Questions 11, 12, 13, 16, 18, 19, 20)
No ○ Go to Question 15. Yes ○ See page 9 of the guide. Print the totals here.
Total partnership tax credits **14A** $
Total partnership income—if a loss, put a minus sign in the last box **14B** $

15. Did the LTC receive any **income from another LTC**? (Exclude any income/losses received you have included at Questions 11, 12, 13, 16, 18, 19, 20)
No ○ Go to Question 16. Yes ○ See pages 10 and 11 of the guide. Print the totals here.
Total LTC tax credits **15A** $
Total active LTC income—if a loss, put a minus sign in the last box **15B** $
Non-allowable deductions **15C** $
Adjusted LTC income (add Boxes 15B and 15C)—if a loss, put a minus sign in the last box **15D** $

16. Did the partnership or LTC receive any **income from overseas**?
No ○ Go to Question 17. Yes ○ See pages 11 to 13 of the guide. Print the totals here.
Total overseas tax paid **16A** $
Total overseas income—if a loss, put a minus sign in the last box **16B** $

17. Did the partnership or LTC receive **income from business activities**?
No ○ Go to Question 18. Yes ○ See page 14 of the guide. Print the total here.
Net income from business activities—if a loss, put a minus sign in the last box **17B** $

18. Did the partnership or LTC receive **income from rental activities**?
No ○ Go to Question 19. Yes ○ See pages 14 and 15 of the guide. Print the total here.
Net income from rental activities—if a loss, put a minus sign in the last box **18B** $

19. Did the partnership or LTC receive any **other income?**

No ◯ Go to Question 20. Yes ◯ See pages 15 to 18 of the guide. Print the total here.

Name of payer

Type of income

Total other income—if a loss, put a minus sign in the last box

19B $

20. Can the partnership claim a **loss from a loss attributing qualifying company (LAQC)?**

No ◯ Go to Question 21. Yes ◯ See page 18 of the guide. Print the amount here.

LAQC loss

20B $ –

21. Total income/loss

Add Boxes 10B to 14B, **15D**, 16B to 19B, then subtract any loss shown in Box 20B. Print the total in Box 21.

Total income—if a loss, put a minus sign in the last box

21 $

22. Can the partnership or LTC claim expenses?

No ◯ Go to Question 23. Yes ◯ See page 18 of the guide. Print the total here.

Total expenses

22 $

23. Total income/loss after expenses

Subtract the amount in Box 22 from the total income in Box 21. Print your answer in Box 23.
This amount must be completely distributed to the partners or owner(s).

Total income after expenses—if a loss, put a minus sign in the last box

23 $

24. Did the partnership or LTC have any losses extinguished on transitioning from a qualifying company (QC) or LAQC?

No ◯ Go to Question 25. Yes ◯ See pages 19 and 20 of the guide. Print the amounts here.

Extinguished losses

24 $

Prior years' deductions

24A $

Deductions claimed this year

24B $

Distribution of income/losses

See page 21 of the guide before distributing income/losses.

25. Distribution of income/loss details attached are for (please tick one):

25 ◯ a partnership (IR 7P) ◯ a look-through company (IR 7L)

Disclosure

26. If the partnership or LTC calculates CFC or FIF income under Question 16, please read page 34 of the guide.
Tick "yes" if additional disclosure is required.

26 No ◯ Go to Question 27. Yes ◯ Go to Question 27.

Declaration

27. Read this declaration and sign the return.

This is a true and correct return for the year ended 31 March 2012.

Signature / / Date

There are penalties for not putting in a tax return or putting in a false return.
Keep a copy of this return for your own records.

28. What to do next

- Remember—the last day for posting us the return is 7 July 2012—read the notes on page 4 for more information.
- Attach all necessary papers to the top of page 3 and make sure the partnership's or LTC's name and IRD number are on all papers.
- All partners and owners must include their share of the partnership or LTC income/losses in their individual tax returns.
- Post the return to us in the envelope supplied, or send to:
  Inland Revenue, PO Box 39090, Wellington Mail Centre, Lower Hutt 5045.

Privacy To find out what may happen to the information you provide on this form, see page 34 of the guide.

OFFICE USE ONLY 29 Operator code ◯ Corresp. indicator ◯ Payment attached ◯ Return cat. ◯ AI Yes ◯ No ◯ Scanner

Inland Revenue
Te Tari Taake

Partnership income/loss distribution

IR 7P
March 2012
2012

- **Read pages 21 to 25 of the IR7 guide before completing this form.**
- **Complete this form and attach it to the top of page 3 of the partnership's IR 7 income tax return.**

Partnership name

IRD number (8 digit numbers start in the second box. 1 2 3 4 5 6 7 8)

Partner's name

IRD number 25A

Share of income/losses

Interest – if a loss, put a minus sign in the last box
25B $

Dividends
25C $

Māori authority distributions
25D $

Overseas income – if a loss, put a minus sign in the last box
25E $

Rental income – if a loss, put a minus sign in the last box
25F $

Other passive income – if a loss, put a minus sign in the last box
25G $

All other income (not already included at Boxes 25B to 25G) – if a loss, put a minus sign in the last box
25H $

LAQC loss
25I $ –

Total income (sum of Boxes 25B to 25I) – if a loss, put a minus sign in the last box
25J $

Proportion of profits/losses (see page 22 of the guide)
25K %

Deduction for extinguished losses (see page 22 of the guide)
$

Share of tax credits

Overseas tax paid
25M $

Imputation credits
25N $

Other tax credits
25O $

Partner's name

IRD number 25A

Share of income/losses

Interest – if a loss, put a minus sign in the last box
25B $

Dividends
25C $

Māori authority distributions
25D $

Overseas income – if a loss, put a minus sign in the last box
25E $

Rental income – if a loss, put a minus sign in the last box
25F $

Other passive income – if a loss, put a minus sign in the last box
25G $

All other income (not already included at Boxes 25B to 25G) – if a loss, put a minus sign in the last box
25H $

LAQC loss
25I $ –

Total income (sum of Boxes 25B to 25I) – if a loss, put a minus sign in the last box
25J $

Proportion of profits/losses (see page 22 of the guide)
25K %

Deduction for extinguished losses (see page 22 of the guide)
$

Share of tax credits

Overseas tax paid
25M $

Imputation credits
25N $

Other tax credits
25O $

Inland Revenue
Te Tari Taake

Individual tax return

Injury Prevention, Rehabilitation, and Compensation Act 2001 • Income Tax Act 2007 • Tax Administration Act 1994

IR 3 2012

1 April 2011 to 31 March 2012

You can file this return online at www.ird.govt.nz
Please see pages 5 and 6 in the guide to see if you need to complete this return.

1 If your IRD number is **not** shown above, print it in Box 1. **1**
(8 digit numbers start in the second box 1 2 3 4 5 6 7 8)

2 If your correct name is **not** shown above, print it in full in Box 2. **2** Mr Mrs Miss Ms
Please put first names **above** and surname **below**

3 If your correct **postal** address is **not** shown above, print the full address in Box 3.
Don't print your tax agent's address here. See page 7 in the guide. **3**
Please put street address or PO Box number **above** and suburb, box lobby or RD and town or city **below**

4 If your correct **street** address is **not** shown below, print it in full in Box 4. **4**
Please put street address **above** and suburb, city, RD or region **below**

5 If your correct date of birth is **not** shown below, print it in Box 5. **5** Day Month Year

6 If your business industry classification (BIC) code is not shown below or has changed, print it in Box 6. **6**

7 If your correct daytime phone number is **not** shown below, print it in Box 7. **7** Prefix Phone number

8 If your correct bank account number is **not** shown below, print it in Box 8. **8** Bank Branch Account number Suffix
For more information about direct crediting, see page 8 of the guide.

9. **Working for Families Tax Credits customers (WfFTC) – adjustments to your family income**
Please check the enclosed *Adjusting your income for Working for Families Tax Credits (IR 215)* form. If you have any adjustments to make and have not told us about them please tick 9A and complete the enclosed IR 215 form.
If you have told us about your adjustments, you don't need to complete an IR 215 form.
Note: If you're not already registered for WfFTC and think you may be entitled to it, read page 8 of the guide. **9A**

10 You may be a non-resident of New Zealand for tax purposes and may need to complete an IR 3NR return instead.
See page 9 in the guide.

Income

Note: If any total is a loss, put a minus sign in the last box provided.

11 Did you receive **family tax credit** from **Work and Income**? Don't include any payments from Inland Revenue.
No Go to Question 11A Yes Enter the amount from your summary of earnings below.
Total family tax credit from Work and Income **11** $

11A Did you receive **income with tax deducted**, as shown on your summary of earnings? See page 11 in the guide.
No Go to Question 12 Yes Copy the amounts from your summary of earnings to the boxes below.
Total PAYE deducted **11A** $
Total gross income **11B** $
ACC earners' levy **11D** $
Total **income not liable** for ACC earners' levy **11C** $
Total tax deducted **11E** $
See page 13 in the guide for a list of income not liable for ACC earners' levy (eg, NZ Super, income-tested benefits, student allowances).

12 Did you receive income from **schedular payments**, as shown on your summary of earnings? See page 14 in the guide.
No Go to Question 13 Yes Copy the schedular payment totals from your summary of earnings to the boxes below.
Total tax deducted **12A** $
Total gross payments **12B** $

13 Did you have any **New Zealand interest** paid or credited to you? Include any interest from partnerships, look-through companies and/or trusts. Keep your interest statements or certificates.
No Go to Question 14 Yes See page 15 in the guide. Print the totals here.
Total RWT **13A** $
Total gross interest **13B** $
If any of your interest was received from a partnership, look-through company, trust and/or estate – please tick 13C. **13C**

14 Did you have any **New Zealand dividends** paid or credited to you or did you receive shares instead of dividends? Include any dividends from partnerships, look-through companies and/or trusts. Keep your dividend statements
No Go to Question 15 Yes See page 17 in the guide. Print the totals here.
Total dividend imputation credits **14** $
Total dividend RWT and payments for foreign dividends **14A** $
Total gross dividends **14B** $
If any of your dividends were received from a partnership, look-through company, trust and/or estate – please tick 14C. **14C**

15 Did you receive any taxable **Māori authority distributions**? Keep your distribution statement
No Go to Question 16 Yes See page 19 in the guide. Print the totals here.
Total Māori authority credits **15A** $
Total Māori authority distributions **15B** $

16 Did you receive any **New Zealand estate or trust income?** See page 19 in the guide.
No Go to Question 17 Yes Print the totals here. Exclude interest, dividends and distributions at Questions 13, 14 or 15, as appropriate.
Total tax paid by trustees **16A** $
Total estate or trust income (excluding amount in Box 16C) **16B** $
Total taxable distributions from non-complying trusts **16C** $

17 Did you receive any **overseas income?** Staple proof of overseas tax paid and a letter detailing any overseas losses to the top of page 3.
No Go to Question 18 Yes See page 20 in the guide. Print the totals here.
Total overseas tax paid **17A** $
Total overseas income **17B** $

Attach your summary of earnings showing any amendments, your Working for Families Tax Credits form and any other information required, to the **top of this page.**

18 Did you receive any **partnership income**? Exclude income at Questions 13, 14, 15, 17, 22 or 24 as appropriate.

No ○ Go to Question 19 Yes ○ See page 26 in the guide. Print the totals here.

Total partnership tax credits

18A $

Total active partnership income

18B $

19 Did you receive any look-through company (LTC) income? Exclude any income/losses received at Questions 13, 14, 15, 17, 22 , 24 and 25 as appropriate.

No ○ Go to Question 20 Yes ○ See page 26 in the guide. Print the totals here.

Total LTC tax credits

19A $

Total active LTC income
If this amount is a loss, put a minus sign in the last box.

19B $

Non-allowable deductions

19C $

Adjusted LTC income (add Boxes 19B and 19C) – if a loss, put a minus sign in the last box.

19D $

20 **Tax credit and income subtotal**

Add the blue Boxes 11E, 12A, 13A, 14A, 15A, 16A, 18A and 19A. Print the total in Box 20A.
Tax credit subtotal

20A $

Add the dark red Boxes 11B, 12B, 13B, 14B, 15B, 16B, 16C, 17B, 18B and 19D. Print the total in Box 20B.
Income subtotal

20B $

21 Did you receive a **shareholder-employee salary** with **no** tax deducted?

No ○ Go to Question 22 Yes ○ See page 27 in the guide. Print the totals here.

Total shareholder-employee salary

21 $

If you **did not receive** a shareholder-employee salary but may in the future – please tick 21A.

21A ○

22 Did you receive any **rents**?

No ○ Go to Question 23 Yes ○ See page 28 in the guide. Print the totals here.

Net rents

22 $

23 Did you receive income from **self-employment**? Don't include any income from your summary of earnings here.

No ○ Go to Question 24 Yes ○ See page 28 in the guide. Print the totals here.

Self-employed net income

23 $

24 Did you receive any **other income**?

No ○ Go to Question 25 Yes ○ See page 30 in the guide. Print the totals here.

Please put name of payer **above**, and type of income **below**

Total other net income

24 $

25 Are you claiming a loss from a **loss attributing qualifying company (LAQC)?**

No ○ Go to Question 26 Yes ○ See page 33 in the guide. Print the totals here.

Amount of loss

25 $ –

26 **Total income**

Add Boxes 20B, 21, 22, 23, 24, and subtract any loss claimed in Box 25. Print your answer in Box 26.

Total income

26 $

27 Are you claiming **expenses** against your income? **Note**: If you've claimed expenses somewhere else in this return, **don't show them here**.

No ○ Go to Question 28 Yes ○ See page 33 in the guide. Print the totals here.

If you paid someone to complete your return, print that person's name in the panel below.

Please put first names **above**, and surname **below**

Total expenses claimed

27 $

28 **Income after expenses**

Subtract Box 27 from Box 26. Print your answer in Box 28.
Use this amount to work out your tax credits.

Income after expenses

28 $

29 Are you claiming **net losses brought forward**?

No ○ Go to Question 30 Yes ○ See page 34 in the guide. Print the net loss amounts in Boxes 29A and 29B.

Amount brought forward

29A $ –

Amount claimed this year

29B $ –

30 **Your taxable income**

Subtract Box 29B from Box 28. Print your answer in Box 30.

Taxable income

30 $

Tax credits

Claim tax credits for donations, childcare or housekeeper payments on the *Tax credit claim form (IR 526).* Don't send in donation receipts with this IR 3 return. See page 35 in the guide.

31 Is your **income at Question 28 under $9,880** and did you earn it by working **20 hours or more a week** or did you receive a sickness benefit, accident compensation payments or earner-related compensation (see page 35 in the guide)?

No ○ Go to Question 32 Yes ○

- To work out if you can claim this tax credit, see page 36 in the guide.
- Copy the number of weeks from Box 4 on page 36 of the guide to Box 31A below.

Note: If you don't fill in Box 31A, we won't be able to calculate your tax credit.

Print the number of weeks here

31A

Print your tax credit here.

31 $

32 **Tax credit for children:** Were you under 15 or under 19 and still at school, at any time from 1 April 2011 to 31 March 2012?
Note: If all your income is interest, dividends, and/or Māori authority distributions, you can't claim this tax credit.

No ○ Go to Question 33 Yes ○ To work out if you can claim this tax credit, see page 37 in the guide.

Print your tax credit here.

32 $

33 If your **income at Question 28 is between $24,000 and $48,000**, you may be able to claim the independent earner tax credit (IETC). See page 38 of the guide to see if you're eligible.

33A **Are you eligible for this tax credit?** To have the IETC included in your assessment you must tick "Yes" below.

33A Yes ○ No ○

33B If you received **excluded overseas income**, complete the dates you received this below.

33B Start (Day Month Year) End (Day Month Year)

Print the number of qualifying months here.

33C

Use the worksheet on pages 39 & 40 in the guide to calculate your IETC.

Print your tax credit here.

33 $

34 Do you have **excess imputation credits brought forward?**

No ○ Go to Question 35 Yes ○ See page 40 of the guide. Print the total here.

34 $ –

Tax calculation

35 Please use the tax calculation worksheet on page 43 of the guide to work out the amount of tax to pay or amount to be refunded.

Tax on taxable income

Transfer the **tax on taxable income** from Box 2 of the worksheet to Box 35. **35** $

Residual income tax

Transfer the **residual income tax** from Box 12 of the worksheet to Box 35A. **35A** $

(Tick one) Credit ○ Debit ○

Tax calculation result

Transfer the **tax calculation result** from Box 14 of the worksheet to Box 35B. **35B** $

(Tick one) Credit ○ Debit ○

36 Are you entitled to claim an **early payment discount**? See page 46 in the guide. **36** Yes ○ No ○

Refunds and/or transfers

Note: If you or your spouse or partner (or ex-spouse or ex-partner) received any Working for Families Tax Credits during the year, this isn't your final refund or tax to pay. You'll need to confirm the details on the Working for Families Tax Credits form and, if there are any changes, staple it to the top of page 3 of this return.

37 Getting a refund — how do you want it paid?

Copy your refund from Box 35B to Box 37. **37** $

Print any overpayment of 2013 provisional tax you would like refunded in Box 37A. **37A** $

Add Box 37 and Box 37A and print your answer in Box 37B. **37B** $

Do you want your refund transferred:

○ to 2013 provisional tax? Print the amount in Box 37C. **37C** $

○ to this year's student loan? Print the amount in Box 37D. **37D** $

○ to someone else's income tax account? Are you "associated"? See page 49 in the guide. **37E** Yes ○ No ○

Name of taxpayer receiving your refund — Their IRD number **37F**

Year ended 31 March — Amount **37G** $

○ to someone else's 2012 student loan? Are you "associated"? See page 49 in the guide. **37H** Yes ○ No ○

Name of taxpayer receiving your refund — Their IRD number **37I**

Amount **37J** $

○ Other — If you want your:
- refund transferred to another tax type, or
- overpaid 2013 provisional tax transferred to pay your 2012 residual income tax debt

staple a note to the top of the front page of this return.

Subtract Boxes 37C, 37D, 37G and 37J from Box 37B. Print your answer in Box 37K. **37K** $

How would you like the balance paid?

- To your bank or other deposit account. Please check your correct account number is shown at Question 8.
- **By cheque** Only tick Box 37L if your bank account details are shown at Question 8 and you now want the balance paid by cheque. **37L** ○

2013 provisional tax

38 Is the amount in Box 35A a debit of more than $2,500?

No ○ Go to Question 39 Yes ○ You may have to pay 2013 provisional tax. See page 50 of the guide, then print the details below.

- Print the option used (S, E or R) in Box 38A. **38A**
- Print your 2013 provisional tax payment in Box 38B. **38B** $ .00

39 Did you at any time during the income year hold rights in a foreign company, unit trust, superannuation scheme or life insurance policy for which disclosure is required?

39 No ○ Go to Question 40 Yes ○ See page 52 of the guide.

40 Is this return for a part-year?

40 No ○ Go to Question 41 Yes ○ See page 53 in the guide. If "Yes", tick the situation below that applies to you.

40A
- ○ You arrived in (or returned to) New Zealand and you're now a tax resident
- ○ You left New Zealand permanently (for more than 325 days)
- ○ You were declared bankrupt
- ○ This return is for a deceased person to the date of their death
- ○ Your balance date changed during the year.

Print the start and end dates the return is for in the spaces provided. **40B** From Day Month Year to **40C** Day Month Year

Notice of assessment and declaration

41 Please see page 53 in the guide, then read and sign the following:

The information in this return is true and correct and represents my assessment for the year ended 31 March 2012 as required under the Tax Administration Act 1994. It is also a correct statement of my earnings for the purposes of the Injury Prevention, Rehabilitation, and Compensation Act 2001.

Signature — Date / /

Please make a copy of this return for your own records.

There are penalties for not filing a tax return or for filing a false return. See page 6 in the guide.

42 What to do next

- Attach your summary of earnings or Working for Families Tax Credits form to your return if you've made any changes on them.
- Print your name and IRD number on all other papers attached to your return.
- Send us your return on or by 7 July 2012, unless you have an extension of time or a non-standard balance date.
- Send your completed return in the envelope supplied, or to the address shown below. Keep a copy for your records.
- If you have tax to pay, you need to pay it by 7 February 2013 to avoid any penalties (if you have a tax agent see page 55 in the guide). You can pay earlier if you want to. If you want to pay now, staple your cheque to the front page of this return.

Privacy

To find out what may happen to the information you provide on this form, see page 62 in the guide.

Injury Prevention, Rehabilitation, and Compensation Act 2001

See page 61 in the guide to understand how the information you provide on this form will be used for ACC purposes.

Send this form to:

Inland Revenue
PO Box 39090
Wellington Mail Centre
Lower Hutt 5045

OFFICE USE ONLY **43** Operator code ○ Corresp indicator ○ Payment attached ○ Return cat ○ AI Yes ○ No ○ Scanner

April 2012

Recent Motu Working Papers

All papers in the Motu Working Paper Series are available on our website www.motu.org.nz, or by contacting us on info@motu.org.nz or +64 4 939 4250.

15-16 Thirkettle, Matt and Suzi Kerr. 2015. “Predicting harvestability of existing *Pinus radiata* stands: 2013-2030 projections of stumpage profits from pre-90 and post-89 forests”

15-15 Fabling, Richard and David C. Maré. 2015. “Production function estimation using New Zealand’s Longitudinal Business Database.”

15-14 Grimes, Arthur, Robert MacCulloch and Fraser McKay. 2015. “Indigenous Belief in a Just World: New Zealand Maori and other Ethnicities Compared.”

15-13 Apatov, Eyal, Richard Fabling, Adam Jaffe, Michele Morris and Matt Thirkettle. 2015. “Agricultural Productivity in New Zealand: First estimates from the Longitudinal Business Database.”

15-12:Laws, Athene, Jason Gush, Victoria Larsen and Adam B Jaffe. 2015. “The effect of public funding on research output: The New Zealand Marsden Fund.”

15-11 Dorner, Zachary and Suzi Kerr. 2015. “Methane and Metrics: From global climate policy to the NZ farm.”

15-10 Grimes, Arthur and Marc Reinhardt. 2015. “Relative Income and Subjective Wellbeing: Intra-national and Inter-national Comparisons by Settlement and Country Type”

15-09 Grimes, Arthur and Sean Hyland. 2015. “A New Cross-Country Measure of Material Wellbeing and Inequality: Methodology, Construction and Results.”

15-08 Jaffe, Adam and Trinh Le. 2015. “The impact of R&D subsidy of innovation: a study of New Zealand firms.”

15-07 Duhon, Madeline, Hugh McDonald and Suzi Kerr. 2015 “Nitrogen Trading in Lake Taupo: An Analysis and Evaluation of an Innovative Water Management Policy.

15-06 Allan, Corey, Suzi Kerr and Campbell Will. 2015. “Are we turning a brighter shade of green? The relationship between household characteristics and greenhouse gas emissions from consumption in New Zealand” (forthcoming)

15-05 Fabling, Richard and Lynda Sanderson. 2015. "Exchange rate fluctuations and the margins of exports”

15-04 Fabling, Richard, Richard Kneller and Lynda Sanderson. 2015. "The impact of tax changes on the short-run investment behaviour of New Zealand firms"

15-03 Sin, Isabelle, Steven Stillman. 2015. “Economic Liberalisation and the Mobility of Minority Groups: Evidence for Māori in New Zealand”

15-02 Grimes, Arthur, Ian Mitchell. 2015. “Impacts of Planning Rules, Regulations, Uncertainty and Delay on Residential Property Development”

15-01 De Rassenfosse, Gaétan, and Adam B. Jaffe. 2015. “Are Patent Fees Effective at Weeding Out Low-Quality Patents?”

14-15 Sin, Isabelle, Richard Fabling, Adam Jaffe, David C. Maré and Lynda Sanderson. 2014. “Exporting, Innovation and the Role of Immigrants.”

14-14 McLeod, Keith, Richard Fabling and David C. Maré. 2014. “Hiring New Ideas: International Migration and Firm Innovation in New Zealand.”