Te Kāwanatanga o Aotearoa

Te Puni Kōkiri
MINISTRY OF MĀORI DEVELOPMENT

Te Matapaeroa Financial Year 2023

Methodology Report

Published by

Te Puni Kōkiri
ISBN: 978-0-478-34561-2
Pipiri/June 2025

Authors

Rajas Kulkarni, Senior Analyst, Te Puni Kōkiri

Angus Prain, Senior Analyst, Te Puni Kōkiri

Melanie Berg, Senior Advisor, Te Puni Kōkiri

Any inquiries, please contact Monitoring & Evaluation team

Disclaimer

Access to the data used in this study was provided by Stats NZ under conditions designed to give effect to the security and confidentiality provisions of the Data and Statistics Act 2022. The results presented in this study are the work of the author, not Stats NZ or individual data suppliers.
These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) and Longitudinal Business Database (LBD) which are carefully managed by Stats NZ.
The results are based in part on tax data supplied by Inland Revenue to Stats NZ under the Tax Administration Act 1994 for statistical purposes. 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.

Contents

Section 1. Introduction

Te Matapaeroa FY 2023 is the fourth iteration of a series about Māori enterprise, published by Te Puni Kōkiri. This report provides information about the data sources, methodology and the approach used to produce data for Te Matapaeroa.

For Te Matapaeroa FY 2023, we have made some changes to the methodology for identifying individual businesses. These methodology changes have resulted in counts, proportions, and trends over time that are different to those reported in Te Matapaeroa 2021. Furthermore, we have added a significant new component to the analysis focussed on employee characteristics. This new material is intended as a replacement and expansion of the 'Significant Employers of Māori' component from previous iterations.

Section 2. Data Sources

2.1 Integrated Data Infrastructure and Longitudinal Business Database

The data source for Te Matapaeroa is Stats NZ's Integrated Data Infrastructure (IDI) and Longitudinal Business Database (LBD). We use several core and derived tables from these databases which we will cover in Section 2.3 below.

The IDI is a large research database holding de-identified microdata about people and households in New Zealand. The IDI contains longitudinal data on most individuals in New Zealand spanning education, health, income, tax, and more. Similarly, the LBD is a large research database holding de-identified microdata about businesses. It is a collection of business administrative data – mainly tax data collected by Te Tari Taake | Inland Revenue (IR), and data from various business surveys. More details on the IDI and LBD and how it is being used can be found in the research paper by Jones, et al. (2022).

With these databases, we can link businesses, owners, and employees to determine the characteristics of businesses and their employees. Figure 1. illustrates the types of data sources available in the IDI and LBD, and the link between the two databases through tax data.

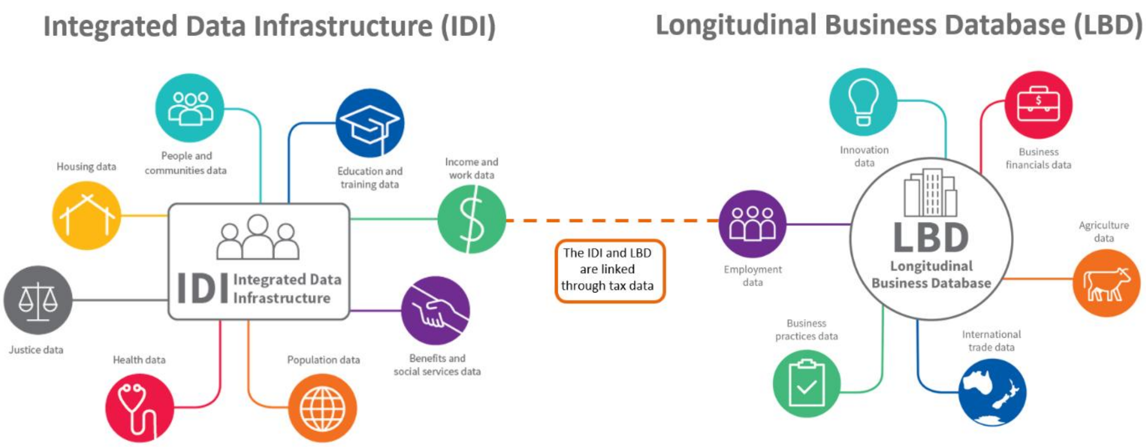

Figure 1: Schematic of the types of data in the IDI and LBD, and the link between the two databases with tax data (source: Stats NZ)

2.2 Productivity and Labour Tables

Created by Richard Fabling and Dave Maré, the Productivity and Labour tables are central to the production of Te Matapaeroa data. These tables provide a stable base business population along with financial and productivity variables such as labour productivity and sales revenue among others. A key advantage of using the Fabling-Maré tables is that businesses are labelled by "Permanent Enterprise Numbers" or "PENTs" which are derived unique identifiers of businesses. These PENTs fix broken[1] 'enterprise numbers' which are unique identifiers for businesses in the LBD environment based on economic activity, location, and ownership for firms. See Fabling (2011), and Fabling and Maré (2019) for more details. The productivity and labour tables allow us to conduct stable longitudinal analysis on Māori-owned businesses.

2.3 Te Matapaeroa Data Sources

Te Matapaeroa 2023 uses tables from the following IDI and LBD sources:

- **Fabling & Maré Labour Tables** – to identify the base business population and their employees, along with other business demographics including employee wages and salaries. These tables include the [pent_year_L_IDI_202401_RFabling], [pent_WP_yr_IDI_202410_RFabling] and [pent_emp_mth_FTE_IDI_202410_RFabling] tables
- **Fabling & Maré Productivity Tables** – to access financial performance measures for businesses.
- **Longitudinal Business Frame (LBF)** – to identify business location of activity and industry of operation.
- **IR's IR7P (Partnership Income Returns), IR4S (Shareholder Income Returns), IR3 (Self-employed Income Returns)** – to identify business owners and sole traders.
- **Stats NZ Derived Personal Details** – to gather ethnicity and gender information on owners and employees.
- **New Zealand Census of Population and Dwellings, 2013 (Census 2013), 2018 (Census 2018) and 2023 (Census 2023) and Department of Internal Affairs (DIA) Birth Records** – to gather Māori descent information for owners and employees.

Inside the Data Lab environment, we have scrips that take these tables as inputs that produce our project-specific derived tables. From these derived tables we produce output-ready summary statistics that are then output through Stats output checkers.

Section 3. Defining Business

A business is an organisation or entity that is actively operating in Aotearoa NZ and is engaged in commercial, industrial, or professional activities. Common business types include Limited Liability Companies,

[1] Enterprise numbers in the LBD can be "broken" by, among other things, changes in the legal status of the firm. This can result in double counting of same businesses or increased rate of entry and exit in the market. It also adds extra noise when conducting longitudinal analysis.

Partnerships, and Sole Proprietors. We classify a 'business' as an entity where the employer and employee(s) are two distinct legal entities. Most of these businesses are Companies[2] and Partnerships[3].

For Te Matapaeroa we only report on businesses that are:

- **Economically Significant**
  A business is economically significant[4] if it meets any one of the following criteria:
  - Annual expenses or sales (subject to GST) of more than $30,000; or
  - 12 month rolling mean employee count of greater than three; or
  - Part of a group of businesses; or
  - Registered for GST and involved in agriculture or forestry; or
  - Over $40,000 of income recorded in the IR10 annual tax return (this includes some units in residential property leasing and rental).
- **Private-for-profit**
  We exclude businesses that are not private-for-profit, including Māori Authorities, iwi-owned businesses, government organisations, embassies, government or locally owned trading entities, and charities.
- **Currently trading ('active')**
  If a business does not have any income, employment, or sales in the financial year of interest, it is deemed inactive.

When identifying **Māori-owned** and **non-Māori-owned businesses** we restrict our business population to:

- **Businesses with at least one identifiable owner actively working in the business ('working proprietor')**
  When business ownership is not explicitly indicated, our population includes businesses without identifiable owners along with businesses owned by other businesses or entities. However, when ownership is explicitly indicated, the baseline population is only those businesses with identifiable owners. In a small proportion of businesses, where the business is partially owned by an individual and partially owned by a trust or business, we only consider the demographic information of the individual registered as the owner of the business to determine the Māori ownership or not of the business.

3.1 Identifying Ethnicity and Descent

To identify ethnicity, we use Stats NZ derived Personal Details Tables, which draw upon multiple data sources to gather ethnicity information on individuals. We use 'total ethnicity' of the owners or employees, which means an individual can belong to multiple ethnic groups. If any one of those ethnic groups is Māori, then that individual is counted as Māori in our data[5].

---

[2] *A company is an entity that is legally separate from the people who own and run the company. It is legislated under the Companies Act 1993.*

[3] *A partnership is when two or more people (partners) form a business, with a partnership agreement setting out how they will share profits, debts, and work. A Partnership is legislated under the Partnership Law Act 2019.*

[4] *See Economically significant enterprise*

[5] *The same owner/employee can also be identified as, for example, 'European' if the individual reported as being of both Māori and European ethnicity.*

We use a combination of DIA birth records, Census 2013, Census 2018 and Census 2023 to gather Māori descent information on individuals. If an individual has said they are of Māori descent in any of these sources, we assign a descent identifier to that person[6].

The ethnicity and descent information on individuals is fixed by how it appears in the most recent records for the period of analysis i.e., 2004 - 2023. For example, if someone identified as Māori descent in Census 2013, that demographic information applies to that individual for all years of our analysis, even prior to 2023 – any business that they might own in, say, 2007 will be classified as Māori-owned.

3.2 Māori-Owned Businesses

Māori-owned businesses are identified using the ethnicity and descent of the working proprietors (WP) - the owners of the business who actively[7] work in the business.

- **Our Primary Māori-owned Business Definition:** A business that is not a Sole Proprietor where at least one WP is of Māori ethnicity or descent and receives a non-zero ownership income in a financial reporting year. This definition includes businesses where a WP may have identified with one or more other ethnicities, such as Pasifika, European, Asian, or Middle Eastern Latin American African (MELAA), or may have no identified ethnicity at all, but were of Māori descent in one or more data source (See Section 3.1).

Where no further qualifications are explicitly stated, this what is meant by Māori owned business in Te Matapaeroa.

3.3 Other business ownership types

While the above definition serves as our primary definition for Māori-owned business, we have also produced data summaries for a wider family of definitions. Below is a summary of some of the variations and ownership features we have included in our data:

- **50% ownership** – A business where 50% of ownership income is earned by individuals of Māori ethnicity or descent in a financial reporting year.
- **Ethnicity only** – A business with a working proprietor of Māori ethnicity, regardless of whether that person identifies as being of Māori descent.
- **Māori sole proprietor** – Otherwise known as a Māori sole trader.
- **Māori working proprietor or Māori sole proprietor** – A business which is either a Māori owned business by our primary definition or is a Māori sole proprietor.
- **Māori-owned business in all active years** – A business which has been captured by our definition (or variant of) as being Māori-owned in all active years of trading. This variation is an attempt to produce an even more longitudinally stable business ownership classification.

[6] Following internal consultation, a choice was made to combine 'yes' responses associated with Māori descent from both Censuses rather than only use one Census year. This may overinflate Māori descent numbers since people's responses to this question may vary over time for various reasons. See Māori descent (information about this variable and its quality) for more information regarding completeness and imputation used in Census 2018.

[7] We classify owners as 'active' if they are drawing upon ownership income from the business. Currently, there is no data in the LBD and IDI on owners who might be active but do not pay themselves ownership income.

- **Māori-owned business in the majority of its active years** - A business which has been captured by our definition (or variant of) as being Māori-owned in the majority of its active years of trading. This slightly weaker definition allows for a business to have a small number of financial years where it would not be labelled as Māori owned by our primary definition but which for the majority of its active years would be labelled as Māori owned by our primary definition.
- **Wahine Māori Business Definition:** A business where at least one working proprietor is a wahine and of Māori ethnicity or descent and receives a non-zero ownership income in a financial reporting year.
- As a comparison, we also report on all **Wāhine Businesses** where at least one owner is a wahine[8] with any ethnicity and receives non-zero ownership income in a financial reporting year.

3.4 Challenges with the 50% Ownership Definition

Te Matapaeroa 2020 used the 50% ownership criteria (see Section 3.3 above) to identify Māori businesses. However, implementing this definition accurately using data in the IDI and LBD is challenging.

As stated earlier, businesses mainly fall into two categories – companies and partnerships. Companies are owned by shareholders who may or may not be working in the business, i.e., some shareholders can be WPs.

Companies also have directors, who operate the business but may or may not own the business. Both directors and shareholders, if working, receive wages. In their tables, Fabling and Maré have attempted to remove individuals who are directors only, therefore identifying shareholder population.

Shareholders receive wages that are a combination of labour they provide to the company and the proportion of the company that they own. To classify a business as being 50% owned by someone of Māori ethnicity (and/or descent), we need to measure the ownership component only, excluding the labour input. Currently the wage data available in the IDI does not provide a breakdown between wages based on ownership and those based on labour input. Because of these limitations it is not possible to accurately identify whether 50% of ownership income is going to someone of Māori ethnicity (and/or descent) for some companies. This is an issue for companies with more than one WP.

[8] *Wahine here refers to the WP reporting themselves as 'female'.*

Figure 2 shows the distribution of WPs in companies in Financial Year 2021. A majority (62.8% of all companies or 48.3% of all businesses in our population) of the companies have a single WP. A further 34.9% of companies (26.8% of businesses in our population) have 2 WP with only 2.3% of companies (1.8% of all businesses in our population) with 3 or more WP. This means that for 37.2% of companies (28.6% of all businesses in our population), it is not currently possible to accurately identify whether 50% or more of ownership income is going to Māori.

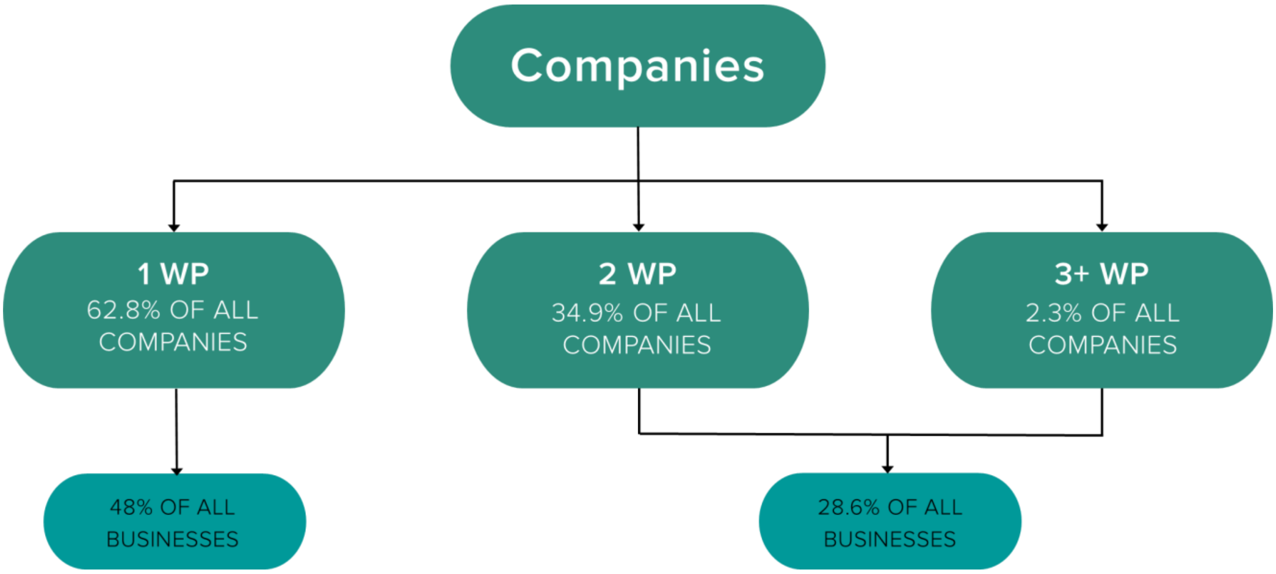

Figure 2: Distribution of working proprietors in companies

In Partnerships, however, partners may or may not be WPs in the business. Partners receive partnership salary, which is based on the proportion of the business they own as set out in the partnership agreement. This means that it is possible to easily identify proportional ownership based on income for these businesses.

3.5 Business financial information

One major update for Te Matapaeroa Financial Year 2023 with respect to previous Te Matapaeroa is the inclusion of financial performance data for Māori owned business.

In addition to the basic count of Māori owned businesses across a range of aggregation variables such and industry and region we also produce, for each aggregation set (e.g. all Māori owned businesses in Northland Region in Agriculture) the following financial summaries:

- Sum of sales
- Median sales per business
- Sum of profits on sales
- Median profit on sales per business
- Sum of zero-rated sales (proxy for exports)
- Median zero-rated sales per business

And draw a direct comparison with non-Māori businesses by way of casting these summaries also as proportions associated with all businesses in the aggregation set.

3.6 Māori employment

Using the Fabling and Maré Labour tables, we can identify all employees that received some employment income in a given calendar month and the amount of remuneration they received. Furthermore, these tables also link employees to the business that employed them.

A connected sequence of months in which an employee is receiving income from a single business is called an employment spell, and an individual's spells can overlap in a given month if they work for multiple employers.

For Te Matapaeroa, spell start months and end months are excluded from employment spells due to peculiarities of wages at the start and end of employment spells. We also explicitly exclude so-called 'short spells', which are employment spells that consist entirely of start and end months. See Fabling, R., & Maré, D. C. (2015) for more details.

Inside the Data Lab environment, we characterise the employment spells per financial year by including only those months of a spell that overlap with a given financial year. For each individual spell's overlap with a financial year, we form the following summaries:

- Sum of wages earned
- Sum of months worked
- Sum of wages earned / sum of months worked ('wage rate' for the spell)

In addition to this, we select a single unique spell for each individual employee per financial year, called the employee's **main job**, which represents the spell that the employee spent the largest number of months in over the course of the financial year. Ties are broken randomly.

For output from the IDI / LBD this summarised 'wage spell' table is summarised further, aggregating over business and employee characteristics, such as business ownership, industry of the business, ethnicity / descent of the employee, employee age bracket, and the geographic location of the employee. These output-ready summaries are what we provide publicly for download as the employee component of the Te Matapaeroa data files. For each aggregation set (e.g. employees ages 20-25 working in Construction in Auckland for a Māori owned business) we calculate:

- Count of unique individuals
- Count of unique employing businesses
- Sum of months worked for all employees
- Sum of wages earned by all employees
- Median monthly wage over all main jobs
- Sum of log wages (sum of the logarithm of the monthly wage rate per spell)
- Count of unique individuals for main jobs

From these summaries one can form representative wage-rates for each aggregation:

- Mean wage per person (total wages / unique people)
- Mean monthly wage (total wages / total months worked)
- Median monthly main job wage (median of the monthly wage rates for all main job spells)

- Geometric mean wage (exponential of (total log wages / total months worked))

Each measure has different characteristics and can be used for different purposes. Due to the presence of outliers in the underlying wages data we recommend using the median monthly wage or the geometric mean wage for stable analysis.

We calculate and include in the data the **wage gap between Māori and non-Māori** based on the median monthly main job wage, noting that using our more comprehensive .csv data files, researchers are free to calculate other wage gaps such as **gender pay gap** using their preferred wage measure from those provided.

Section 4. Key Descriptors for Businesses

The section below provides technical details on the key variables for Māori-owned businesses and sole traders. This information is included to aid in the interpretation of the datasets, and the Te Matapaeroa insights report.

4.1 Location and Industry

Industry and location are two key characteristics of any enterprise (business or sole trader). The Fabling-Maré Labour Tables provide a permanent industry and location for enterprises based on economic activity of the business. For Te Matapaeroa, we want to report on all the locations and industries Māori businesses are operating in so we can track and report on the movement within these. To do this, we use the plant[9] level data from the Longitudinal Business Frame (LBF). The majority of enterprises only have one location and industry, but larger businesses often have more than one plant, which may be undertaking different activities. Te Matapaeroa reports on all the active[10] locations and industries for businesses and sole traders.

We report on Industry using the Australian and New Zealand Standard Industrial Classification (ANZSIC)[11], 2006. We report geographic data using Stats NZ defined Regional Councils and Territorial Authorities, and by Te Puni Kōkiri Region.

4.2 Firm Size

We calculate 'firm size' using full-time equivalent (FTE) total labour input of the employees and the WPs. The FTE data comes directly from Fabling-Maré Labour Tables (Fabling & Maré, 2015).

4.3 Exporter Proxy

We define two classes of exporters – Major exporters and Minor exporters. Using the GST data, we define 'Major exporters' as "a business with at least NZ$ 250,000 of zero-rated[12] GST sales or where a business has 50% of more of total sales as zero-rated sales with a minimum total sale of NZ$ 250,000". We define

[9] *An enterprise comprises one or more geographic units (geo) or 'plants', which are operations undertaken at a single location. Each plant has a unique location and industry based on the activity undertaken.*

[10] *"Active" means that the plant must have some economic activity, unless a business has no active plans, in which case we chose the last active plant to report on industry and location.*

[11] *See ANZSIC*

[12] *https://www.ird.govt.nz/gst/charging-gst/zero-rated-supplies*

'Minor exporter' as those with at least some zero-rated sales but those who do not meet criteria for being a Major exporter.

There are several caveats with using this proxy –

- Not all zero-rated sales will be exports. This means that some businesses can be classified as 'exporter' incorrectly.
- The threshold of NZ$ 250,000 could potentially remove some of the smaller exporters or could incorrectly classify as 'minor exporter' in a particular year of smaller zero-rated GST.
- This definition does not account for consistency in zero-rated sales could result in businesses being classified as exporters in one year (e.g. if they sold assets due to liquidation), and then not in next.

References

Fabling, R. (2011). Keeping it Together: Tracking Firms in New Zealands's Longitudnal Business Database. *Motu Working Paper 11-01*.

Fabling, R., & Maré, D. C. (2015). Addressing the absence of hours information in linked employer-employee data. *Motu Working Paper 15-17*.

Fabling, R., & Maré, D. C. (2019). Improved productivity measurement in New Zealand's Longitudinal Business Database. *Motu Working Paper*.

Jones, C., McDowell, A., Galvin, V., & Adams, D. (2022, April 28). Building on Aotearoa New Zealand's Integrated Data Infrastructure. *Harvard Data Science Review, 4*. Retrieved from https://hdsr.mitpress.mit.edu/pub/9dkr3v8v

Te Puni Kōkiri
MINISTRY OF MĀORI DEVELOPMENT

Te Puni Kōkiri, Te Puni Kōkiri House, 143 Lambton Quay, PO Box 3943, Wellington, New Zealand
PHONE Waea 0800 875 663 (0800 TPK MMD), **FAX** Waea Whakaahua 0800 875 329 (0800 TPK FAX)
WEB tpk.govt.nz, **FACEBOOK** facebook.com/tepunikokiri