

# Queen's Biology Student Funding Compared to Market Basket Measure (MBM) Line

Caelan Johnson

2024-12-19

## Purpose of this Plot

My aim is to create an accurate visualization of how Queen's Biology graduate students' minimum stipends compare to Canada's poverty metric, the Market Basket Measure (MBM), over time.

## Definitions

One of the trickiest parts of creating this visualization with confidence is sorting out the definitions. Below are those that I've identified so far as being important. Please note that the official definitions are long, complicated, and can be full of jargon. I've corresponded with some analysts from Statistics Canada for clarity in some cases, particularly with the square root equivalency and how to appropriately use it.

**Market Basket Measure (MBM):** This is Canada's official poverty metric. It's a monetary value that represents the annual income needed for a reference family of four—two adults and two children.

<https://www12.statcan.gc.ca/census-recensement/2021/ref/dict/az/definition-eng.cfm?ID=pop165>

**Economic family:** "Economic family refers to a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common-law union, adoption or a foster relationship."

<https://www23.statcan.gc.ca/imdb/p3Var.pl?Function=Unit&Id=33863>

**Disposable income:** This is the monetary value that we use to compare to the MBM. It is the total income of an economic family after deducting not only taxes but also "non-discretionary expenses".

<https://www150.statcan.gc.ca/n1/pub/75f0002m/75f0002m2019014-eng.htm>

**Square root equivalency:** Because the MBM is set for a reference family of 4, Statistics Canada uses this square root equivalency method to adjust the value to be applicable to economic families of different sizes. The formula is  $P_i = P_r * (\sqrt{S_i}/2)$ , where  $P_i$  is the poverty line for the economic family in question,  $P_r$  is the poverty line for the reference family, and  $S_i$  is the size of the family in question.

<https://www150.statcan.gc.ca/n1/pub/75f0002m/75f0002m2024004-eng.htm>

**Base year:** Every 8 to 10 years, the MBM goes through a "rebasings" process, where the market basket is redefined based on modern product availability. The "base year" refers to the year that was used to characterise the market basket for the given basing effort. The link below is, if I understand correctly, a document that updates the public on the progress of the 2018 rebasing.

<https://www150.statcan.gc.ca/n1/pub/75f0002m/75f0002m2019013-eng.htm>

# Queen's Biology Stipend and Tuition Data

I've created a dataset that has the stipend and tuition values for Biology graduate students.

## Tuition data

In the GitHub repository you will find folders with the archived Queen's tuition files that I have collected from their website (<https://www.queensu.ca/registrar/tuition-fees/archived>). The “total” values for MSc and PhD students are what I've used as the datapoints for this dataset. I'm putting a focus on domestic students for this visualization since the tuition data for them goes back the furthest.

## Stipend data

I received hard copies of records of the Biology department's graduate stipends from the previous Graduate Assistant, Joanne Surette (thank you Joanne!). I found the minimum stipend available for each year and used those as the datapoints for the stipend amounts, separating MSc from PhD where applicable. I've uploaded these documents as a PDF file to the GitHub in the “Grad Stipends” folder.

For the most recent couple of years, I found the department's stipend information off the website (included below). I'm not aware of the archive of this information being available online, but will update if I learn more.

<https://biology.queensu.ca/current-students/graduate-students/fees-and-funding-0>

## Loading the stipend data

```
stip = read.csv("Stipend Data.csv")
```

## Calculate DI

Because graduate students cannot be enrolled unless they pay tuition and fees, I am treating the tuition and fees as “non-discretionary expenses”, and will subtract these values from the stipends of domestic students in order to find their disposable income.

```
stip$MSc_Domestic_DI = stip$MSc - stip$DomFees  
stip$PhD_Domestic_DI = stip$PhD - stip$DomFees
```

## Standardize to 2023 dollar

In order to make sure we can compare dollar amounts between years, I'll standardize all of these values to the 2023 dollar. I use the priceR package to do this.

```
stip$MSc_Domestic_DI_2023 = afix(price = stip$MSc_Domestic_DI  
  ,from_date = stip$Year  
  ,country = "Canada"  
  ,to_date = 2023  
  )
```

```
## Retrieving countries data
```

```
## Generating URL to request all 296 results
## Retrieving inflation data for CA
## Generating URL to request all 64 results
```

```
stip$PhD_Domestic_DI_2023 = afi(price = stip$PhD_Domestic_DI
  ,from_date = stip$Year
  ,country = "Canada"
  ,to_date = 2023
)
```

```
## Retrieving countries data
```

```
## Generating URL to request all 296 results
## Retrieving inflation data for CA
## Generating URL to request all 64 results
```

## MBM Data

Statistics Canada has an excellent tool you can use to get any MBM data you're looking for. You can find the tool using the link below.

<https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1110006601>

Here is the citation information: Statistics Canada. Table 11-10-0066-01 Market Basket Measure (MBM) thresholds for the reference family by Market Basket Measure region, component and base year

DOI: <https://doi.org/10.25318/1110006601-eng>

I've customized the table to select for years going back to 2006, standardized to the 2023 dollar concept, and selecting the population size that corresponds to Kingston.

```
mbm = read.csv("1110006601_databaseLoadingData.csv")
```

## Selecting a base year

The Statistics Canada database doesn't allow for the most recent base year, 2018, to be used to find an MBM value further than the year 2015. When I reached out to an analyst, they informed me that rather than mixing the use of different base years, I should extrapolate the 2018 base year from 2015 back to 2006, based on the year by year percent change shown in the 2008 base year MBM values. They gave me an example, but I haven't quite been able to make sense of it. Right now, I'll restrict this visualization to having a lower bound for year of 2015.

## Filtering MBM data

I'll filter for the MBM values we're looking for—a base year of 2018. I'll also select for just the columns that have the year and the value of the MBM.

```
mbm = filter(mbm, Base.year == "2018 base")

mbm = select(mbm, REF_DATE, VALUE)
```

## Economic family of size 1

According to the definition of an economic family, students that are rooming together are not considered to be an economic family. This means we need to adjust the MBM values to be for an economic family of size 1 to represent grad students not living with their parents. To do this, we need to use the square root equivalency. We'll use:

$$P_i = P_r * (\sqrt{S_i}/2)$$

Where  $P_i$  is the poverty line for the economic family in question,  $P_r$  is the poverty line for the reference family, and  $S_i$  is the size of the family in question.

Because the size of the economic family we're interested in is 1, we can then write:

$$P_i = P_r * (\sqrt{1}/2)$$

Which simplifies to:

$$P_i = P_r/2$$

So, I'll divide the MBM values by 2.

```
mbm$Econ_fam_1 = mbm$VALUE/2
```

## Creating the plot

### Joining the datasets

Now that we have the stipend and MBM data ready, I'll organize them into a dataset that is ggplot friendly.

```
stipmbm = tibble(year = c(stip$Year, stip$Year,mbm$REF_DATE)
  , value = c(stip$MSc_Domestic_DI_2023
    , stip$PhD_Domestic_DI_2023
    , mbm$Econ_fam_1
  )
  , data_source = c(rep("Domestic MSc DI",length(stip$MSc_Domestic_DI_2023))
    , rep("Domestic PhD DI",length(stip$PhD_Domestic_DI_2023))
    , rep("MBM", length(mbm$Econ_fam_1))
  )
)
```

### Make the plot

```
(ggplot(data = stipmbm, mapping = aes(x = year, y = value, color = data_source))
+ geom_point()
+ geom_line()
+ scale_x_continuous(limits = c(2015,2023), breaks = seq(from = 2015, to = 2023, by = 1))
+ theme_classic()
+ labs(
```

```

x = "Year"
, y = "CAD Standardized to 2023"
, title = "Queen's Biology Domestic Grad Student DI Compared to MBM"
, color = "Line"
)
)

```

```

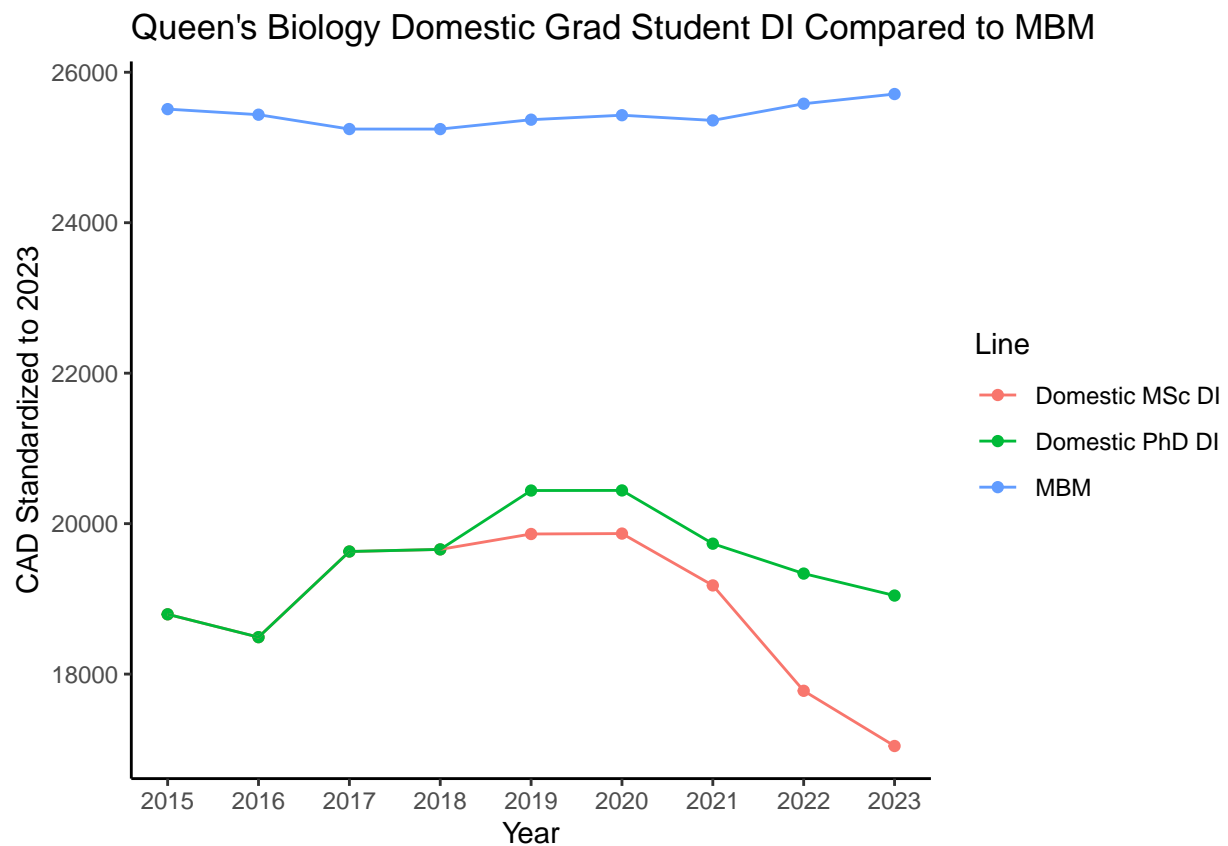
## Warning: Removed 34 rows containing missing values or values outside the scale range
## ('geom_point()').

```

```

## Warning: Removed 34 rows containing missing values or values outside the scale range
## ('geom_line()').

```



## Domestic Student DI Visualization

Here's a plot of just the domestic student disposable income, going back to 2006, corrected to the 2023 dollar.

```

(ggplot(data = filter(stipmbm, data_source != "MBM")
, mapping = aes(x = year
, y = value
, color = data_source
)
)

```

```

    )
+ geom_point()
+ geom_line()
+ scale_x_continuous(limits = c(2006,2023), breaks = seq(from = 2006, to = 2023, by = 1))
+ theme_classic()
+ theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
+ labs(
  x = "Year"
  , y = "CAD Standardized to 2023"
  , title = "Queen's Biology Domestic Grad Student DI"
  , color = "Line"
)
)
)

```

```

## Warning: Removed 16 rows containing missing values or values outside the scale range
## ('geom_point()').

```

```

## Warning: Removed 16 rows containing missing values or values outside the scale range
## ('geom_line()').

```

