Research Examples using QWI Explorer

Local Employment Dynamics

Using the QWI to Study Startup Firms in the Pittsburgh Metro Area

Suppose we are interested in pursuing a research agenda focused on startup firms in the Pittsburgh Metro Area. We can use QWI data, with tabulations by firm age, to study these new firms.

The QWI provides detailed tabulations by 5 firm age categories: 0-1 years, 2-3 years, 4-5 years, 6-10 years, and 11+ years. For this example, let us define "startup" firms as those that are of age 0-1. With this setup, we can investigate a variety of research questions.

We'll use QWI Explorer for our analysis: http://qwiexplorer.ces.census.gov/

Research Question 1: Who Works at Startups?

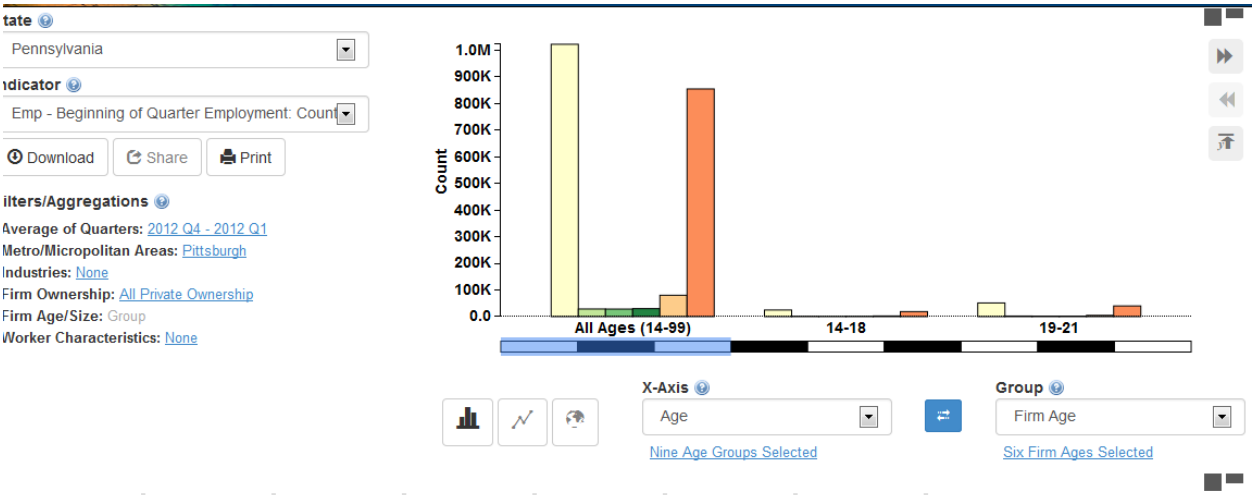
Does the composition of the workforce at startups differ from the workforce at other firms? To answer this, we can compare the distribution of various worker characteristics at startups with the distribution at all firms.

For example, let's look at the age distribution.

In QWI Explorer, follow these steps:

- Set **State** to **Pennsylvania**
- Set **Indicator** to **Emp** (Beginning-of-Quarter Employment)
- Change **X-Axis** to **Worker Age**
    - Click "Continue" on the conflict box (moving Worker Age from Group to X-Axis)
    - Click the blue text under the X-Axis dropdown, which reads "Eight Age Groups Selected," and check the box for "All Ages" (so that all boxes are checked)
- Change **Group** to **Firm Age**
    - Click "continue" on the conflict box (this will change ownership to **Private**)
    - Click "close" on the "No Data Available" box (firm age is not available for the most recent quarter. We'll resolve this by choosing the filter)
    - Click the blue text under the Group dropdown, which reads "Five Firm Ages Selected", and check the box for "All Firm Ages" (so that all boxes are checked)
- Under **Filters**:
    - **Quarters** – click the blue text, deselect 2013Q1, and check all 4 quarters in 2012. This will generate the average employment over these 4 quarters.
    - **Sub-State Geography** – click the blue text, change the dropdown to Metro/Micro Area, and check the box for **Pittsburgh**

The screen should look similar to this:



	All Firm Ages	0-1 Years	2-3 Years	4-5 Years	6-10 Years	11+ Years
All Ages (14-99)	1,020,781	28,798	28,409	30,001	80,124	853,447
14-18	24,831	1,236	1,081	1,131	2,574	18,806
19-21	51,265	2,219	2,006	2,052	4,726	40,260
22-24	65,788	2,626	2,721	2,538	5,771	52,130
25-34	212,942	6,692	7,614	7,683	18,154	172,799
35-44	196,736	5,266	5,403	5,928	15,334	164,803
45-54	238,594	5,616	5,263	5,923	17,237	204,552
55-64	179,876	4,125	3,211	3,630	12,119	156,791
65-99	50,746	1,016	1,107	1,112	4,205	43,304

From looking at the table, we can see that the age distribution at startups differs from all firms – the largest Worker Age category for 0-1 Years is 25-34, while the largest category for All Firm Ages is 45-54.

Let's calculate the percentage of workers that are under age 35, by firm age. Click the Download button, and select "Download Table as Zip." Open the CSV file.

In a new row, add together employment for ages 14-18, 19-21, 22-24, and 25-34 (should be rows 3-6). Do this for the "All Firm Ages" and "0-1 Years" columns.

In the next row, divide this total by the "All Ages" Row. This gives us the share of workers that are age 14-34 in All Firms and in Startup Firms. The spreadsheet should look similar to this:

	All Firm Ages	All Firm Ages Flags	0-1 Years	0-1 Years Flags	2-3 Years	2-3 Years Flags	4-5 Years	4-5 Years Flags	6-10 Years	6-10 Years Flags	11+ Years	11+ Years Flags
All Ages (14-99)	1020781	10	28798	10	28409	10	30001	10	80124	10	853447	10
14-18	24831	10	1236	10	1081	10	1131	10	2574	10	18806	10
19-21	51265	10	2219	10	2006	10	2052	10	4726	10	40260	10
22-24	65788	10	2626	10	2721	10	2538	10	5771	10	52130	10
25-34	212942	10	6692	10	7614	10	7683	10	18154	10	172799	10
35-44	196736	10	5266	10	5403	10	5928	10	15334	10	164803	10
45-54	238594	10	5616	10	5263	10	5923	10	17237	10	204552	10
55-64	179876	10	4125	10	3211	10	3630	10	12119	10	156791	10
65-99	50746	10	1016	10	1107	10	1112	10	4205	10	43304	10
Sum 14-34	354826		12773									
Share 14-34	35%		44%									

This table suggests that startup firms disproportionately employ younger workers: 44% of workers at startups are under age 35, compared to only 35% at all firms.

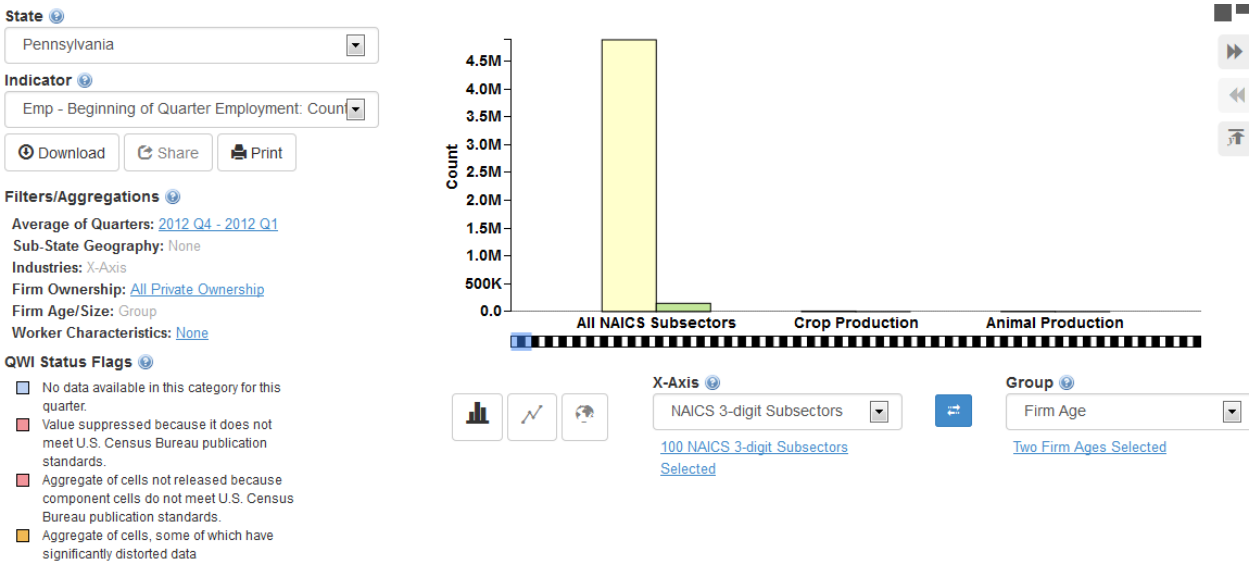
Research Question 2: Which Industries have the Highest Percentage of Employment in Startups?

Let's look at the share of startup employment in 3-digit NAICS Subsectors in Pennsylvania.

In QWI Explorer, follow these steps:

- Set **State** to **Pennsylvania**
- Set **Indicator** to **Emp** (Beginning-of-Quarter Employment)
- Change **X-Axis** to **NAICS 3-digit Subsectors**
    - Click the blue text under the Group dropdown, which reads "99 NAICS 3-digit Subsectors Selected, and check the box for "000 All NAICS" (so that all boxes are checked)
- Change **Group** to **Firm Age**
    - Click "continue" on the conflict box (this will change ownership to **Private**)
    - Click "close" on the "No Data Available" box (firm age is not available for the most recent quarter. We'll resolve this by choosing the filter)
    - Click the blue text under the Group dropdown, which reads "Five Firm Ages Selected", and check ONLY the boxes for "All Firm Ages" and "0-1 Years" (so only 2 boxes checked)
- Under **Filters**:
    - **Quarters** – click the blue text, deselect 2013Q1, and check all 4 quarters in 2012. This will generate the average employment over these 4 quarters.
    - **Note that we can not filter by sub-state geography** – At the NAICS 3 and NAICS 4 level, firm age and firm size are only available at the state level

The screen should look similar to this:



	All Firm Ages	0-1 Years
All NAICS Subsectors	4,878,959	149,668
Crop Production	11,659	382
Animal Production	5,450	245
Forestry and Logging	832	156
Fishing, Hunting and Trapping	60	
Support Activities for Agriculture and Forestry	3,123	164
Oil and Gas Extraction	5,402	198
Mining (except Oil and Gas)	12,472	436
Support Activities for Mining	16,458	1,197
Utilities	23,477	122
Construction of Buildings	51,620	2,561

To find the Subsectors with the highest percentage of startups, we need to move this to a Spreadsheet, calculate percentages, and then sort.

Click the Download button, and select "Download Table as Zip." Open the CSV file.

Perform the following steps:

- In the next empty column (likely column F, depending on your spreadsheet's formatting), type "Share 0-1 Years" in the first row.
- In the second row, divide the value in "0-1 Years" by the value in "All Firm Ages" (You should be able to type = D2/B2, although the column letters may differ depending on your spreadsheet).
- After that formula calculates, select that cell, and then click and drag on the bottom right-hand corner of that cell to copy the formula down.
- Finally, select this "Share 0-1 Years" column, click "Sort and Filter," and then click "Sort largest to smallest"

The results (ignoring the 'Divide by Zero' rows) should look like this:

	All Firm Ages	All Firm Ages Flags	0-1 Years	0-1 Years Flags	Share 0-1 Years
Postal Service	13	10	4	10	31%
Scenic and Sightseeing Transportation	821	10	186	12	23%
Private Households	4063	10	894	10	22%
Forestry and Logging	832	10	156	10	19%
Food Services and Drinking Places	370335	10	29590	10	8%
Social Assistance	149556	10	11079	10	7%
Support Activities for Mining	16458	10	1197	10	7%
Personal and Laundry Services	59391	10	3943	10	7%
Primary Metal Manufacturing	38650	10	2501	10	6%
Petroleum and Coal Products Manufacturing	6368	10	376	11	6%
Accommodation	51417	10	2882	10	6%
Truck Transportation	61159	10	3236	10	5%
Support Activities for Agriculture and Forestry	3123	10	164	10	5%
All NAICS Subsectors	4878959	10	149668	10	3%

Note that we should probably disregard the Postal Service, since its employment levels are so small.

Otherwise, we can note that several industries have particularly high levels of their employment in startups.

Research Question 3: Does Regional Variation in Startup Concentration Predict Employment Growth?

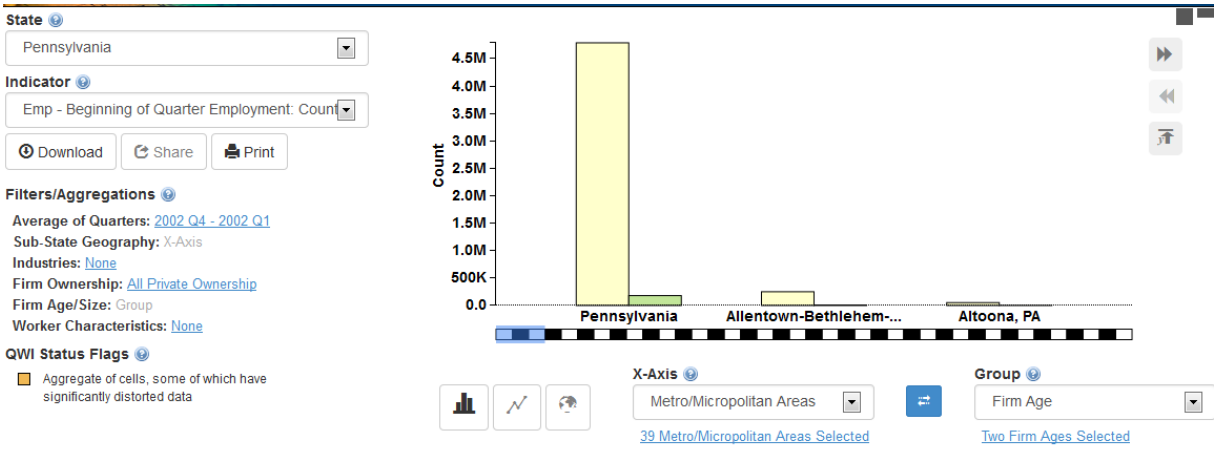
Finally, suppose we want to study the relationship between startup concentration and overall employment growth. For instance, does a large concentration of startups drive employment growth? If so, this would have interesting policy implications – perhaps localities would want to do more to attract and support startups.

To start to answer this question, let's examine regional variation in the concentration of startup employment. Specifically, we will test whether metro areas with the highest concentration of startup employment in 2002 experienced high employment growth over the next decade.

In QWI Explorer, follow these steps:

- Set **State** to **Pennsylvania**
- Set **Indicator** to **Emp** (Beginning-of-Quarter Employment)
- Change **X-Axis** to **Metro/Micro Areas**
  - Click the blue text under the X-Axis dropdown, which reads "Metro/Micro Areas Selected," and check the box for "42 Pennsylvania" (so that all boxes are checked)
- Change **Group** to **Firm Age**
  - Click "continue" on the conflict box (this will change ownership to **Private**)
  - Click "close" on the "No Data Available" box (firm age is not available for the most recent quarter. We'll resolve this by choosing the filter)
  - Click the blue text under the Group dropdown, which reads "Five Firm Ages Selected", and check ONLY the boxes for "All Firm Ages" and "0-1 Years" (so only 2 boxes checked)
- Under **Filters**:
  - **Quarters** – click the blue text, deselect 2013Q1, and check all 4 quarters in 2002. This will generate the average employment over these 4 quarters.

The screen should look like this:



	All Firm Ages	0-1 Years
Pennsylvania	4,786,505	180,005
Allentown-Bethlehem-Easton, PA-NJ (PA part)	251,323	9,380
Altoona, PA	52,396	1,800
Bloomsburg-Berwick, PA	31,995	1,161
Bradford, PA	14,944	457
Chambersburg, PA	41,198	1,553
DuBois, PA	27,832	1,041
East Stroudsburg, PA	39,113	1,884
Erie, PA	105,806	3,948
Gettysburg, PA	26,803	1,120
Harrisburg-Carlisle, PA	252,000	13,761
Huntingdon, PA	9,423	366

Similar to Example 2, we want to identify the Metro Areas with the highest percentage of startups.

Click the Download button, and select "Download Table as Zip." Open the CSV file.

Perform the following steps:

- In the next empty column (likely column F, depending on your spreadsheet's formatting), type "Share 0-1 Years" in the first row.
- In the second row, divide the value in "0-1 Years" by the value in "All Firm Ages" (You should be able to type = D2/B2, although the column letters may differ depending on your spreadsheet).
- After that formula calculates, select that cell, and then click and drag on the bottom right-hand corner of that cell to copy the formula down.
- Finally, select this "Share 0-1 Years" column, click "Sort and Filter," and then click "Sort largest to smallest"

The resulting output should look like this:

	All Firm Ages	All Firm Ages Flags	0-1 Years	0-1 Years Flags	Share 0-1 Years
New York-Northern New Jersey-Long Island, NY-NJ-PA (PA part)	6768	10	525	10	7.8%
Harrisburg-Carlisle, PA	252000	10	13761	10	5.5%
Sunbury, PA	24818	10	1319	12	5.3%
Lewistown, PA	12992	10	644	12	5.0%
Indiana, PA	24312	10	1196	10	4.9%
Oil City, PA	17847	10	876	10	4.9%
East Stroudsburg, PA	39113	10	1884	10	4.8%
Selinsgrove, PA	14556	10	699	12	4.8%
Johnstown, PA	47894	10	2291	10	4.8%
Warren, PA	12758	10	574	10	4.5%

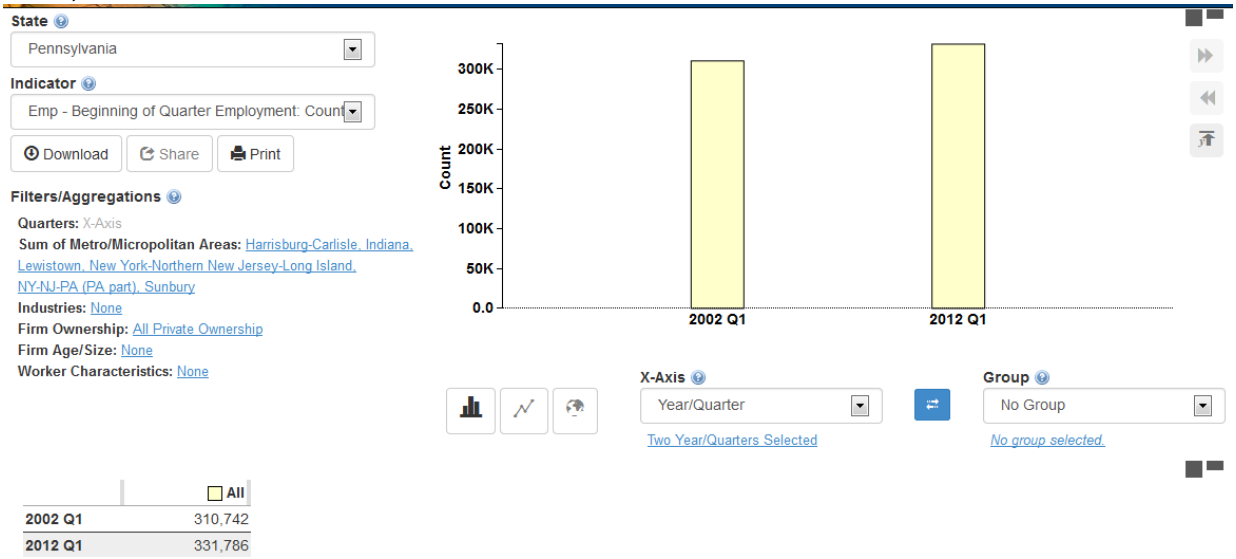
Now, let's focus on the 5 Metro areas with the highest percentage of startup employment. Let's examine how much their employment grew between 2002 and 2012. Here are the steps in QWI Explorer:

In QWI Explorer, follow these steps:

- Set **State** to **Pennsylvania**
- Set **Indicator** to **Emp** (Beginning-of-Quarter Employment)
- Change **X-Axis** to **Year/Quarter**
    - Click the blue text under the X-Axis dropdown, which reads "Eight Year/Quarters Selected," and check ONLY the boxes for 2002Q1 and 2012Q1.
- Change **Group** to **No Group**
- Under **Filters**:
    - **Firm Ownership** – click the blue text, and select the radio button for "All Private Ownership." We want to focus our analysis on private firms only (Since we identified startup concentration by region for only private firms, we want to examine employment growth for private firms – important to be consistent)
    - **Sub-State Geography** – click the blue text, change the dropdown menu to micro/metro area, and select the 5 metro areas that had the highest concentration of startup employment in 2002:
        - Harrisburg-Carlisle
        - Indiana
        - Lewistown
        - New York-Northern New Jersey-Long Island
        - Sunbury

      This will generate the total employment for these 5 metro areas.

So far, the screen should look like this:



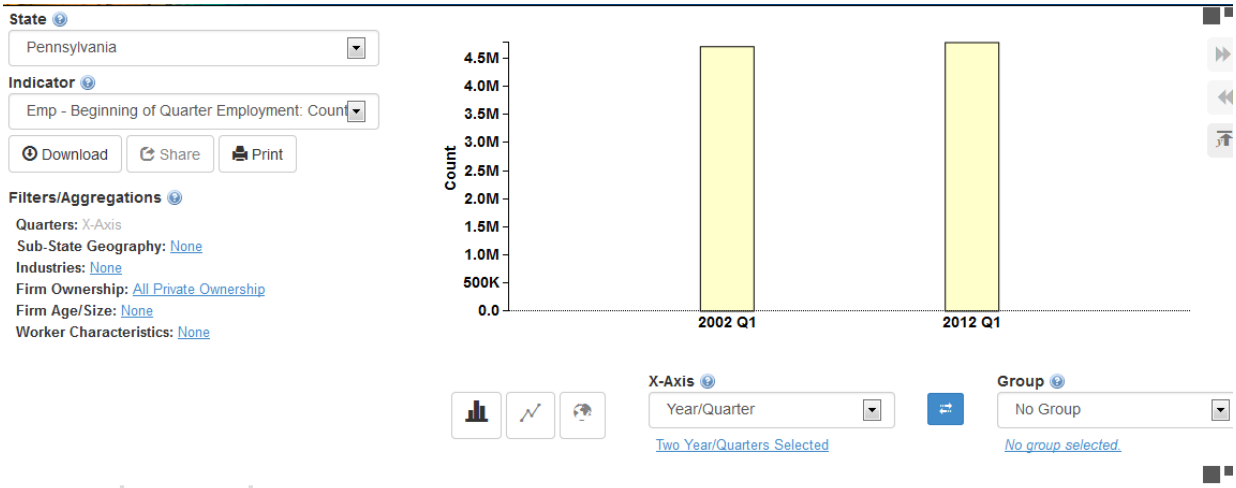
	All
2002 Q1	310,742
2012 Q1	331,786

Employment in these metro areas grew from 310,742 in 2002Q1 to 331,786 in 2012Q1.

To calculate the percentage growth, we can download this to a spreadsheet and divide the two rows. You would find that (331,786 – 310,742) / 310,742 = 6.8%, suggesting that total employment grew by 6.8% in these regions.

How does this compare to the state of Pennsylvania as a whole? Simply remove the filters: click on the blue text next to "Sum of Metro/Micro Areas," and click "Check None."

The screen should now look like this:



	All
2002 Q1	4,703,170
2012 Q1	4,775,714

Notice that employment in the state of Pennsylvania grew from 4,703,170 in 2002Q1 to 4,775,714 in 2012Q1. Calculating the percentage growth gives us (4,775,714 – 4,703,170) / 4,703,170 = 1.5%. This suggests that employment in the 5 metro areas with the highest concentration of startup firms in 2002 grew over four times faster than employment in the entire state (6.8% vs 1.5%)!!

Remember, correlation is not causation – we can't say that the concentration of startups necessarily caused the higher rate of employment growth, only that a correlation exists. Still, this is an interesting result that can motivate future research.