Why the Ratio of all Transit Users to Daily Transit Users is Almost Always Higher for Households than for Individuals.

The likelihood that someone in a household has a rare trait is almost proportional to the number of people in the household, while the likelihood that someone in the household does something fairly common does not rise so fast. Here I will give two simple numerical examples: The case where the propensity among members of a household to use transit is independent among household members, and the case where transit use is correlated but frequency is not. Finally I will close with a few more general relationships.

Assume for simplicity that all households have two members, M and F, and that 10% of the people use transit daily (D), 40% use transit weekly but not daily (W), and 50% do not use transit at all (N).

I. The Case where there is no correlation between how often M and F use transit.

Let's start by assuming there is no correlation between how often M and F use transit.

Out of 100 households, we would expect the following

A. Ten households where M is a daily commuter. So all 10 of those households use transit daily, as shown in the table below, where F uses transit daily 20% of the time, weekly 80% of the time.

Do M and F use transit daily, weekly, or not at all for each combination with a 1% chance of occurring. Group 1: M is a Daily Transit User											
Household Number:	1	2	3	4	5	6	7	8	9	10	
M	D	D	D	D	D	D	D	D	D	D	
F	D	W	W	W	W	N	N	N	N	N	

B. Forty households where M is a weekly transit user, and again, F uses transit daily 10% of the time, weekly 40% of the time, and not at all 50% of the time. As the next table shows, 10% of those households use transit daily, while 90% of those households use transit weekly. Because there are 40 such households, we have 4 more households using transit daily and 36 that use transit weekly.

Do M and F use transit daily, weekly, or not at all for each combination with a 4% chance of occurring. Group 2: M is a Weekly Transit User											
Household Number:	1	2	3	4	5	6	7	8	9	10	
M:	W	W	W	W	W	W	W	W	W	W	
F:	D	W	W	W	W	N	N	N	N	N	

C. Fifty households where M does not use transit, and again, F uses transit daily 10% of the time, weekly 40% of the time, and not at all 50% of the time. As the next table shows, 10% of those households use transit daily, and 40% of those households use transit weekly. Because there are 50 such households, we have 5 more households using transit daily and 20 that use transit weekly.

Do M and F use transit daily, weekly, or not at all for each combination with a 5% chance of occurring. Group 3: M does not use transit.											
Household Number:	1	2	3	4	5	6	7	8	9	10	
M:	N	N	N	N	N	N	N	N	N	N	
F:	D	W	W	W	W	N	N	N	N	N	

Adding all three distributions, we have 19 households using transit daily, and 56 who use it weekly, for a total of 75 households that use transit. The proportion of *households* with someone who uses transit daily is almost twice (19%/10%) the proportion of *individuals* who use transit daily. The proportion of households who use transit is only 1.5 times the proportion of individuals who use transit (75%/50%).

Thus, among individuals, the ratio of transit users to daily users is 5:1; but among households the ratio is 75:19, a little bit less than 4:1.

II. The case where there is perfect correlation between *whether* M and F use transit, but not correlation between how often they use transit.

We might also consider the case where there is a perfect correlation between H and W regarding whether they use transit at all, but no correlation on whether they use transit weekly or daily. One might expect this situation, for example, if non-transit users live in places where there is no transit. In such a case, 50% of households use no transit, and of the remaining 50% of households, 20% of the people use transit daily, while 80% use it weekly.

Out of 100 households, we would expect the following

A. Ten households where M is a daily commuter. All 10 of those households use transit daily, as shown in the table below, where F uses transit daily 20% of the time and weekly 80% of the time.

Do M and F use transit daily, weekly, or not at all for each combination with a 1% chance of occurring. Group 1: M is a Daily Transit User											
Household Number:	1	2	3	4	5	6	7	8	9	10	
M	D	D	D	D	D	D	D	D	D	D	
F	D	D	W	W	W	W	W	W	W	W	

B. Forty households where M is a weekly transit user. Again, F uses transit daily for 20% of the households, and weekly 80%. As the next table shows, 20% of those households use transit daily, while 80% of those households use transit weekly. Because there are 40 such households, we have 8 more households using transit daily and 32 that use transit weekly.

Do M and F use transit daily, weekly, or not at all for each combination with a 4% chance of occurring. Group 2: M is a Weekly Transit User											
Household Number:	1	2	3	4	5	6	7	8	9	10	
M:	W	W	W	W	W	W	W	W	W	W	
F:	D	D	W	W	W	W	W	W	W	W	

C. Fifty households where M does not use transit, and given the perfect correlation, F does not use transit either. So none of these households use transit daily or weekly.

Do M and F use transit daily, weekly, or not at all for each combination with a 5% chance of occurring. Group 3: M does not use transit.											
Household Number:	1	2	3	4	5	6	7	8	9	10	
M:	N	N	N	N	N	N	N	N	N	N	
F:	N	N	N	N	N	N	N	N	N	N	

Adding all three distributions, we have 18 households using transit daily, and 32 who use it weekly, for a total of 50 households that use transit.

Adding all three distributions, we have 19 households using transit daily, and 56 who use it weekly, for a total of 75 households that use transit. The proportion of *households* with someone who uses transit daily is almost twice (18%/10%) the proportion of *individuals* who use transit daily. Given our assumption, the proportion of households who use transit is the same as the proportion of individuals who use transit.

Thus, among individuals, the ratio of transit users to daily users is 5:1; but among households the ratio is 50:18, approximately 2.75. That is fairly similar to the 17.4:6.6 ratio among households nationwide, as reported by the Census Bureau. While that does not mean that the ratio of transit users to daily users is actually 5:1 either nationwide or in Maryland, this example illustrates why one would expect that the ratio of transit users to daily users could be much higher among individuals than among households.

If there was a negative correlation between propensities to use transit, then the ratio might actually be less for individuals than households. In a community with fair transit where people are limited to owning one car, for example, that might occur. But there is no reason to think that this is what happens in Maryland.

III. The Ratio as a function of probability

Going back to our example where transit use is independent, if the probability that someone uses transit is *p*, then the probability that neither people in a two person household use transit is $(1\text{-}p)^2$, which means that the probability that someone uses transit is $1\text{-}(1\text{-}p)^2$, which equals $2p\text{-}p^2$. So if the probability that someone is a daily user is p_d, then the probability that someone in the household is a daily user is $2p_d - p_d^2$. So the ratio of household transit users to household daily transit users, in general, will be:

$$\text{Ratio} = \frac{2p - p^2}{2p_d - p_d^2}$$

Here is a table with some example values, using this formula:

p_d: Proportion of people using transit daily	*P*: Proportion of people who use transit	Ratio of individual transit users to daily users	Ratio of household transit users to daily users
0.15	0.05	3	2.85
0.3	0.1	3	2.68
0.45	0.15	3	2.51
0.6	0.2	3	2.33
0.2	0.05	4	3.69
0.4	0.1	4	3.37
0.6	0.15	4	3.03