The Economics Honors Thesis of Mike Filicicchia

Part I: Exploring Internal and External Factors of Child Development (Professor Frank Stafford)

Part II: The Determinants of Consumer Sentiment in the Housing Market (Professor Robert Barsky)

Mike Filicicchia

Part I: Exploring Internal and External Factors of Child Development

Introduction

A child's development is a complex mixture of nature and nurture, internal and external, belief and practice. I will take and in-depth look at measures of a child's psychological health as well as socially constructive activity. I weigh external factors such as the frequency of positive parent-child interaction, parental disciplinary techniques and philosophies, and positive parental behavioral modeling (parenting by example). In addition, I weigh internal factors such as a child's personal faith in order to account for ways in which a child is not merely the product of his/her household environment.

Methods

I worked exclusively with data from the 2002 Child Development Supplement since I found that much of the 1997 data from the CDS seemed to be incorrectly labeled. Also, I was able to use data from the "child file" in 2002, which allowed me to compare a child's responses and own feelings of worth against their parent's assessment of it. The 2002 data also included helpful summary statistics, such as Behavior Problem Index (BPI) scores and a parental warmth index. I used data from the time diaries in addition to the main portion of the CDS in order to accurately measure child participation in specific activities.

Introduction to Index Variables

Here I will introduce the five main index variables I used in my analysis. I first analyzed a statistic labeled "parental warmth" (WARMTH02), which mostly includes measures of how often a parent talked with their child about their interests, participated with them in their favorite activities, and gave them verbal affirmation. In many ways, I think this measure would have been better labeled "positive parental involvement", because it did not include any sort of disciplinary behavior, which might seem "cold" to children. For example, a father who was very involved in his children's lives and spoke very lovingly to them on a regular basis would score extraordinarily high on this index even if he was an alcoholic who beat his wife and children in fits of rage. Another issue with this measure is that, though it is a mean of seven other measures (each ranging from 1-5), its distribution is not quite normal (mean = 3.90, std. dev = 0.66), so there is some trouble distinguishing between observations at the high end.

My second crucial index variable is a measure of a child's "positive behavior" (POSBEH02). I was much more pleased with the naming of this variable, as(in my opinion) it truly measured what I would deem "positive behavior" such as helping siblings, demonstrating patience, doing diligent work, and obeying parents. Like parental warmth, this statistic is a mean over many (ten) other measures and was constrained to an interval of 1-5, but has the undesirable property of being strongly skewed left (mean = 4.13, std. dev = 0.60)

The last three index variables of interest are all related to "behavior problems", and so they are labeled the Behavior Problem Index (BPI). There were separate measures for externalizing (aggressive) behaviors such as cheating, arguing, disobedience, restlessness, stubbornness, temper, etc.—17 in total and internalizing (withdrawn or sad) behaviors such as fearfulness, confusion, regret, worthlessness, unhappiness, worry, etc.—14 in total, labeled BPI_E02 and BPI_N02, respectively. The combination of the two was known as the "total BPI", labeled BPI_T02. BPI_E02 has values ranging from 0 to 17 and has a fairly desirable distribution (mean = 5.75, std. dev = 4.20). Similarly, BPI_N02 ranges from 0 to 14 and but is significantly less symmetric (mean = 3.16, std. dev = 3.16).

Parental Warmth and Child Behavior

The very first regression I ran sought to relate these five key index variables. Specifically, I wanted to see how the one parental measure (WARMTH02) related to the three measures of child behavior (POSBEH02, BPI_E02, and BPI_N02), so I regressed the parental behavior on the child behaviors to seek how much explanatory power a child's behavior had on their parent's interactions with them. I first dropped all observations where the information was "not ascertained". What I found was quite surprising:

```
      Source |       SS       df       MS              Number of obs =    2871
-------------+------------------------------           F(  3,  2867) =   32.68
       Model |  40.9607158     3  13.6535719           Prob > F      =  0.0000
    Residual |  1197.68664  2867  .417749089           R-squared     =  0.0331
-------------+------------------------------           Adj R-squared =  0.0321
       Total |  1238.64735  2870  .431584444           Root MSE      =  .64634

------------------------------------------------------------------------------
    WARMTH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    POSBEH02 |     .19033   .0261343     7.28   0.000     .1390861    .2415738
     BPI_E02 |  -.0124257   .0041931    -2.96   0.003    -.0206475   -.0042039
     BPI_N02 |   .0220062   .0052517     4.19   0.000     .0117087    .0323038
       _cons |   3.110539   .1230384    25.28   0.000     2.869286    3.351791
------------------------------------------------------------------------------
```

Here we see two very intuitive and (I daresay) expected results from the data. First, it is quite clear that positive parental involvement in a child's life is highly correlated with the child's display of socially constructive behaviors (there is an incredibly strong positive relationship between WARMTH02 and POSBEH02). Similarly, positive parental involvement is highly correlated with a decrease in aggressive outward behavior (described by the incredibly strong negative relationship between WARMTH02 and BPI_E02). However, I certainly did not foresee the last result on the list and have some trouble describing it. There seems to be a (very) significant increase in withdrawn or sad behavior in children related to a parent's warmth and positive involvement in their life once we control for all the outward signals of psychological well-being. It seems (initially) that if parents are heavily involved in their children's lives, their children are likely to be "good citizens" in all the outward measures, but also internally ridden with all sorts of negative feelings. Since this result is so surprising to me, I seek out the culprit for this effect in the next section.

Parental Warmth and Psychological Strife

First, I wanted to see what the relationship was between these two variables (WARMTH02 and BPI_N02) without controlling for anything else. My result was fairly expected:

```
      Source |       SS       df       MS              Number of obs =    2867
-------------+------------------------------           F(  1,  2865) =    4.57
       Model |  45.6567827     1  45.6567827           Prob > F      =  0.0326
    Residual |  28625.2166  2865   9.9913496           R-squared     =  0.0016
-------------+------------------------------           Adj R-squared =  0.0012
       Total |  28670.8734  2866  10.0037939           Root MSE      =  3.1609

------------------------------------------------------------------------------
     BPI_N02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    WARMTH02 |  -.1920122   .0898231    -2.14   0.033    -.3681367   -.0158877
       _cons |   3.908818   .3548692    11.01   0.000     3.212994    4.604643
------------------------------------------------------------------------------
```

On the whole, increased positive parental involvement in a child's life predicts a decrease in internal psychological strife in that child. This result (if conducting a one-sided t-test) is significant at the 2% level. However, I was still curious what it could be about parental warmth that caused the increase in psychological strife when controlling for outward behavior variables. In order to assess the underlying causes for my unexpected result, I broke down parental warmth into each of its seven separate components and regressed our measure of inward psychological strife (BPI_N02) against all of them to see if I could pinpoint a culprit or two.

I will quickly identify the variables (A-G):
A: Frequency with which a parent tells a child they love him/her
B: Frequency with which a parent spends time with the child doing a favorite activity
C: Frequency with which a parent talks to their child about the child's interests
D: Frequency with which a parent tells their child they are appreciated
E: Frequency with which a parent talks to their child about the child's relationships
F: Frequency with which a parent talks to their child about the news
G: Frequency with which a parent talks to their child about the child's day

The results are once again quite remarkable:

```
      Source |       SS       df       MS              Number of obs =    2867
-------------+------------------------------           F(  7,  2859) =    3.72
       Model |  258.606237     7  36.9437481           Prob > F      =  0.0005
    Residual |  28412.2672  2859  9.93783391           R-squared     =  0.0090
-------------+------------------------------           Adj R-squared =  0.0066
       Total |  28670.8734  2866  10.0037939           Root MSE      =  3.1524

------------------------------------------------------------------------------
     BPI_N02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
```

Q21E13A	.1458155	.0740698	1.97	0.049	.0005799	.291051
Q21E13B	-.1383701	.0645968	-2.14	0.032	-.2650311	-.0117091
Q21E13C	-.1455126	.0805905	-1.81	0.071	-.3035339	.0125087
Q21E13D	-.1075953	.0809564	-1.33	0.184	-.2663342	.0511436
Q21E13E	.200767	.06436	3.12	0.002	.0745703	.3269636
Q21E13F	-.0643489	.0552622	-1.16	0.244	-.1727068	.044009
Q21E13G	-.0881534	.0916133	-0.96	0.336	-.2677882	.0914815
_cons	3.849187	.412114	9.34	0.000	3.041116	4.657257

Needless to say, we have identified our two culprits. Holding constant most measures of a parent's involvement in a child's life, it's clear that the more they told the child they loved them and the more they talked to their child about his/her relationships, the more likely the child was to have sad or withdrawn tendencies. What's even more surprising is that these effects are still slightly detectable even when we DON'T control for other forms of a parent's positive involvement in their child's life:

BPI_N02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21E13A	.0116046	.0634435	0.18	0.855	-.1127949	.1360041
_cons	3.108734	.2906887	10.69	0.000	2.538754	3.678715

BPI_N02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21E13E	.0514407	.053641	0.96	0.338	-.0537382	.1566197
_cons	2.973316	.2042283	14.56	0.000	2.572866	3.373765

Here we see that telling your child you love them or talking about relationships has absolutely no relationship to internal BPI (and possibly even causes more problems). This of course makes no sense unless we consider that it may do harm as often as it does good. Clearly we have identified two possible red flags in parent-child interaction. My best interpretation for this data is that parents who spend an especially high fraction of their verbal interaction with their children telling them they love them and inquiring about their relationships, the more psychologically unhealthy their child is likely to be. Similarly, it seems that parents who simply ask their children about their interests and then participate with them in those interests, are more likely to have psychologically healthy children. I don't claim any sort of causation here. It could be that children who "need" to be told they're loved or talk with their parents about their relationships tend to be especially unstable children to begin with, but it is also possible that engaging in these sorts of discussions at an unusually high rate might even create these sorts of insecurities/distresses in a child's life, or perhaps just indicate an overly nosey parent who is likely to foster these sorts of fears in their children.

Introducing the Internal BPI

In this section and the next, I seek to isolate exactly why these two activities: telling a child you love them and talking with them about their relationships seem to increase this BPI measure. In order to

do so, I break apart the internal BPI into its 14 component parts and analyze their relationship with each of these activities. All factors in the BPI start with the prefix "Q21B29", and their suffixes are described here:

B: Feels or complains that no one loves him/her
C: Is rather high strung, tense, and nervous
E: Is too fearful or anxious
H: Is easily confused, seems to be in a fog
L: Has trouble getting along with other people his/her age
N: Feels worthless or inferior
O: Is not liked by other people his/her age
P: Has a lot of difficulty getting his/her mind off certain thoughts
T: Is unhappy, sad or depressed
U: Is withdrawn, does not get involved with others
X: Cries too much
Z: Is too dependent on others
AA: Feels others are out to get him/her
DD: Worries too much

Telling a Child You Love Them

Since I'm running a regression on 14 variables, I can expect that at least one of them is going to be significant at the 10% level, really not making it significant at all. Therefore, when looking at the regression, I will work at the 1% level to ensure that the factors I'm considering truly are influencing the data.

Q21E13A	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21B29B	-.0104502	.0396291	-0.26	0.792	-.0881548	.0672543
Q21B29C	.0109155	.0338005	0.32	0.747	-.0553605	.0771914
Q21B29E	.0333739	.0343804	0.97	0.332	-.034039	.1007868
Q21B29H	-.0073904	.0391078	-0.19	0.850	-.0840728	.069292
Q21B29L	-.0341398	.0491067	-0.70	0.487	-.1304281	.0621486
Q21B29N	-.0061177	.0537814	-0.11	0.909	-.111572	.0993366
Q21B29O	-.0192698	.0555826	-0.35	0.729	-.1282559	.0897163
Q21B29P	.0061077	.033871	0.18	0.857	-.0603064	.0725218
Q21B29T	-.1390991	.049385	-2.82	0.005	-.2359331	-.0422651
Q21B29U	-.061223	.0513591	-1.19	0.233	-.1619277	.0394817
Q21B29X	.2011896	.0414223	4.86	0.000	.119969	.2824102
Q21B29Z	-.0592906	.0377486	-1.57	0.116	-.133308	.0147268
Q21B29AA	.0209653	.0593878	0.35	0.724	-.0954821	.1374127
Q21B29DD	.1092838	.0337757	3.24	0.001	.0430566	.1755109
_cons	4.403701	.0821596	53.60	0.000	4.242603	4.564799

Clearly, there are three standout factors at the 1% level. Interestingly, one of them even represents a downward movement in BPI. It seems that telling your child you love them is associated with a decrease

in sadness/depression, but interestingly enough, is also correlated with children crying and worrying too much (these latter two are strong enough to outstrip all of the other variables, most of which seem to have some sort of downward effect on BPI). Of course, the contradiction is almost obvious because we normally associate crying with sadness, but in this case it might be that parents whose children cry a lot will often tell them that they love them, even when the crying has nothing to do with the child being sad. On the other hand, the statistic about worry is a little harder to explain. It doesn't seem natural that a parent's response to their child's worry is to tell them that they love them. It could be that if a parent consistently tells their child that they're loved, the child may wonder if they're so unlovable that they must be constantly told they are loved. These statistical measures cannot show the sincerity of the parents' words, and it's quite possible that the parents who tell their children that they love them all the time do not do so with much heartfelt emotion, and therefore the child worries that maybe it is all just a cover-up for their actual lack of love.

Talking with a Child About Relationships

Q21E13E	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21B29B	.0922483	.0470001	1.96	0.050	.0000908	.1844059
Q21B29C	-.0197836	.0400874	-0.49	0.622	-.0983869	.0588197
Q21B29E	.0285382	.0407751	0.70	0.484	-.0514135	.1084899
Q21B29H	-.0470375	.0463818	-1.01	0.311	-.1379828	.0439078
Q21B29L	.0257215	.0582406	0.44	0.659	-.0884764	.1399194
Q21B29N	-.0494335	.0637847	-0.78	0.438	-.1745022	.0756353
Q21B29O	-.0242728	.0659209	-0.37	0.713	-.1535303	.1049847
Q21B29P	.0595638	.040171	1.48	0.138	-.0192033	.1383309
Q21B29T	-.0575898	.0585707	-0.98	0.326	-.1724349	.0572553
Q21B29U	-.1879284	.0609119	-3.09	0.002	-.3073642	-.0684927
Q21B29X	.0675679	.0491268	1.38	0.169	-.0287597	.1638956
Q21B29Z	-.0014608	.0447699	-0.03	0.974	-.0892455	.0863238
Q21B29AA	.0492521	.070434	0.70	0.484	-.0888545	.1873587
Q21B29DD	.1064637	.0400579	2.66	0.008	.0279182	.1850091
_cons	3.542637	.0974413	36.36	0.000	3.351574	3.733699

Once again I will operate at the 1% significance level. The most glaring observation is the 'U' variable that basically states that the more withdrawn a child is, the less likely their parents are to talk with them about their relationships. Regarding causation, this could truly go either way. Either a child doesn't have any friends and therefore his/her parents have nothing to talk about regarding relationships, or it could be that if child relationships are stressed in conversation with parents, then that child is more likely to pursue friendships and have increased social interaction. Again, the upward pressure seems to be from worry here, though a case could be made for not feeling loved (the 'B' variable), mental preoccupation (the 'P' variable), or crying too much (the 'X' variable). This finding is not surprising since a large proportion of a child's worries will be regarding his/her relationships. Again, it could be that a worrisome child will bring these matters up often in conversation, or that a parent who overly stresses talking about a child's relationships could cause them to overanalyze their interactions with others and begin to worry about them.

The Importance of Religion

Beyond what a parent might do to influence their child's good or bad behavior, it is worth noting that children are not merely the product of the parenting styles they were subject to. Many children have internal reasons for their good behavior that lie entirely outside of any way they were brought up, and here I seek to analyze the influence of religion and spirituality in that behavior. It's important to note here that all of the following religious variables apply to children ages 12 and above. I will quickly label the variables I used regarding religion:

Q23J3: Importance of religion to child
Q23J3A: Child's comfort from religion
Q23J3B: Child's religious attendance
Q23J3C: Child's participation in religious clubs/activities

I wanted to look at the relationship between these variables and a child's positive behavior scores as well as BPI scores. The results:

```
      Source |       SS       df       MS              Number of obs =    1005
-------------+------------------------------           F(  4,  1000) =   10.12
       Model |  13.5905044     4   3.3976261           Prob > F      =  0.0000
    Residual |  335.851104  1000  .335851104           R-squared     =  0.0389
-------------+------------------------------           Adj R-squared =  0.0350
       Total |  349.441608  1004  .348049411           Root MSE      =  .57953

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       Q23J3 |   .1255798   .0335352     3.74   0.000     .0597723    .1913873
      Q23J3A |   .0258692   .0294696     0.88   0.380      -.03196    .0836985
      Q23J3B |   -.017483   .0145348    -1.20   0.229    -.0460052    .0110392
      Q23J3C |   .0145262   .0129966     1.12   0.264    -.0109775    .0400299
       _cons |   3.694238   .0818846    45.12   0.000     3.533553    3.854923
------------------------------------------------------------------------------

------------------------------------------------------------------------------
     BPI_E02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       Q23J3 |  -.8548594   .2384255    -3.59   0.000    -1.322731   -.3869877
      Q23J3A |  -.0015398   .2095197    -0.01   0.994    -.4126886     .409609
      Q23J3B |   .1312279   .1033379     1.27   0.204    -.0715562    .3340119
      Q23J3C |  -.1745852   .0924018    -1.89   0.059    -.3559089    .0067385
       _cons |   8.160012   .5821748    14.02   0.000     7.017587    9.302436
------------------------------------------------------------------------------

------------------------------------------------------------------------------
     BPI_N02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       Q23J3 |  -.4305436   .1870415    -2.30   0.022    -.7975824   -.0635047
      Q23J3A |   .0160918   .1643653     0.10   0.922    -.3064487    .3386323
      Q23J3B |    .008917   .0810672     0.11   0.912    -.1501642    .1679983
      Q23J3C |  -.0972312   .0724879    -1.34   0.180    -.2394772    .0450147
       _cons |   4.629931    .456708    10.14   0.000     3.733715    5.526147
------------------------------------------------------------------------------
```

These results were almost exactly as I had predicted. It's clear that the most important factor (religiously) regarding a child's good behavior is how important their religion is to them. The more important a child's religion, the more likely they are to model good behavior and the less likely they are to exhibit unhealthy psychological feelings. Also, once we take religious importance into account, it seems that the comfort a child receives from their religion doesn't have much of an effect at all on these measures. However, as I predicted, religious attendance actually has negative effects on both of the outward measures once we control for the other three variables. Clearly, taking a child to a religious service, if it is not important to them, can have some damaging effects. This is an important point that I think most parents probably miss. If religious attendance results in children finding religion more important, it could end up working for the best regarding their behavior, but if this is not the keep, it seems harmful for them to keep attending. Also, as I expected, it seems that in addition to finding religion important, it is especially helpful if children are involved in outside religious activities, most likely because it gets them the personal attention and sense of community they need in order to feel better about themselves and also because it gives them a group of individuals to participate in constructive behavior with.

Disciplinary Methods

Six Years and Under

I began by analyzing disciplinary methods for children less than 6 years of age. First I wanted to examine a set of questions that asked parents how they would respond if their child had a temper tantrum. A "1" score was assigned to a "yes" response, and a score of "5" was assigned to a "no" response. Therefore, positive coefficients in this category actually demonstrate a negative relationship (since the higher the score, the less likely a parent is to incorporate this method of discipline). However, it is worth noting here that the question asked in the survey was entirely hypothetical and not necessarily a reflection of the sorts of disciplinary methods actually enforced in the home, though I take them as a trustworthy representation. There was some data regarding actual parental disciplinary practices, but I opted against using that data because it asked how many times a given disciplinary method was incorporated in the last week, making most observations equal to zero, thereby giving low variation and, more importantly, an inability to separate lack of disciplinary action with non-response. The variables are labeled according to this method:

Q21F7A: Ground the child	B: Spank the child	C: Talk to the child
E: Ignore it	F: Send child to his/her room	I: Give child "time out"
J: Hug/console child	K: Hit/threaten child	

```
      Source |       SS       df       MS              Number of obs =      48
-------------+------------------------------           F(  7,    40) =    2.52
       Model |  3.18640435     7  .455200621           Prob > F      =  0.0302
    Residual |  7.22359643    40  .180589911           R-squared     =  0.3061
-------------+------------------------------           Adj R-squared =  0.1847
       Total |  10.4100008    47  .221489378           Root MSE      =  .42496

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
```

Q21F7A	(dropped)					
Q21F7B	-.0264743	.0389234	-0.68	0.500	-.1051415	.0521929
Q21F7C	.0199134	.0461859	0.43	0.669	-.0734318	.1132586
Q21F7E	.1053505	.0489937	2.15	0.038	.0063306	.2043703
Q21F7F	.1064838	.035434	3.01	0.005	.034869	.1780985
Q21F7I	-.0601259	.0342667	-1.75	0.087	-.1293813	.0091296
Q21F7J	.0525824	.0456318	1.15	0.256	-.039643	.1448078
Q21F7K	.0025617	.0821339	0.03	0.975	-.1634371	.1685605
_cons	3.588228	.4825756	7.44	0.000	2.612907	4.56355

There are really only three observations of significance here. The first is to be expected: ignoring the problem is correlated with lower positive child behavior. However, the other two significant observations seem almost contradictory. On the one hand, it appears that giving a child time out at this age is correlated with increased positive behavior, while children who would be sent to their room tend not to be very helpful, and this is significant beyond the 1% level, statistically a better indicator of poor behavior than ignoring the problem! Notably, these coefficients are quite high for the types of variables we're looking at. A "yes" response is a downward movement of 4 points in each of these explanatory variables, meaning that a "yes" response to either "ignore the problem" or "send child to his/her room" lowered the positive behavior score by .42 points on average (the equivalent of 7/10 of a standard deviation in positive behavior scores). This simultaneous finding is odd to me because it's hard to really delineate between the two forms of discipline. It's quite paradoxical that the two most similar forms of discipline have the most opposite effects. My best inference is that sending a child to their room may not be any form of discipline at all. Many children enjoy their rooms and have had it customized to their liking, so it could be that sending a child to their room is essentially like ignoring it in the sense that they are not really punished (and perhaps even rewarded) for their outburst. However, giving a child time out not only forces a sort of exclusion, but it also requires a child to reflect on what they have just done. My findings were quite similar when I regressed the BPI scores on these variables (time-out was good, sending a child to their room and ignoring them were bad). However, the only statistically significant result from those regressions was a positive relationship between giving a child time-out and their decrease in internal BPI. For now, I will endorse this form of discipline above the others for this youngest age group.

<u>6-9 Years of Age</u>

We now move on to the 6-9 age group. The same question was asked to parents of children in this age group regarding how they would respond if their child threw a sort of temper tantrum or used profanity in a fit of anger. However, for this age group, I included a few more explanatory variables that seemed age-appropriate.

Q21G17A: Ground the child	B: Spank the child	C: Talk with the child
D: Give the child chores	E: Ignore it	F: Send child to room
G: Take away allowance	H: Take away privileges	I: Give time-out
J: Hug or console child	K: Hit or threaten child	O: Other punishment

```
      Source |       SS       df       MS              Number of obs =     872
-------------+------------------------------           F( 12,   859) =    3.55
       Model |  12.8761542    12  1.07301285           Prob > F      =  0.0000
    Residual |  259.416316   859  .301998039           R-squared     =  0.0473
-------------+------------------------------           Adj R-squared =  0.0340
       Total |   272.29247   871  .312620516           Root MSE      =  .54954

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21G17A |   .0403214   .0111749     3.61   0.000     .0183882    .0622546
     Q21G17B |   -.002543   .0109832    -0.23   0.817    -.0241001    .0190142
     Q21G17C |   .0063934    .010813     0.59   0.554    -.0148297    .0276164
     Q21G17D |  -.0130294   .0173222    -0.75   0.452    -.0470282    .0209693
     Q21G17E |   .0641033   .0159133     4.03   0.000     .0328699    .0953367
     Q21G17F |   .0161314   .0097439     1.66   0.098    -.0029933    .0352561
     Q21G17G |   -.037413   .0176504    -2.12   0.034     -.072056   -.0027701
     Q21G17H |   .0129439   .0101774     1.27   0.204    -.0070317    .0329194
     Q21G17I |    .005863   .0105348     0.56   0.578    -.0148139    .0265399
     Q21G17J |  -.0380327   .0527602    -0.72   0.471    -.1415867    .0655214
     Q21G17K |  -.1531498   .1378813    -1.11   0.267    -.4237736    .1174739
     Q21G17O |  -.1031498   .1378813    -0.75   0.455    -.3737735     .167474
       _cons |   5.295207   1.035216     5.12   0.000     3.263359    7.327056
------------------------------------------------------------------------------

      Source |       SS       df       MS              Number of obs =     872
-------------+------------------------------           F( 12,   859) =    4.70
       Model |  917.548006    12  76.4623339           Prob > F      =  0.0000
    Residual |  13983.5242   859  16.2788408           R-squared     =  0.0616
-------------+------------------------------           Adj R-squared =  0.0485
       Total |  14901.0722   871  17.1080049           Root MSE      =  4.0347

------------------------------------------------------------------------------
     BPI_E02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21G17A |  -.3483888    .082045    -4.25   0.000    -.5094209   -.1873566
     Q21G17B |  -.1020421    .080638    -1.27   0.206    -.2603127    .0562286
     Q21G17C |   -.002783   .0793884    -0.04   0.972     -.158601    .1530349
     Q21G17D |  -.0631579   .1271779    -0.50   0.620    -.3127738    .1864579
     Q21G17E |   -.571369   .1168338    -4.89   0.000    -.8006822   -.3420558
     Q21G17F |  -.0200052   .0715392    -0.28   0.780    -.1604172    .1204068
     Q21G17G |   .0478725   .1295877     0.37   0.712    -.2064731    .3022181
     Q21G17H |  -.0765495   .0747219    -1.02   0.306    -.2232084    .0701094
     Q21G17I |  -.1030389   .0773453    -1.33   0.183    -.2548468     .048769
     Q21G17J |  -.2785946   .3873614    -0.72   0.472     -1.03888     .481691
     Q21G17K |   .4702127   1.012314     0.46   0.642    -1.516685    2.457111
     Q21G17O |   .9702127   1.012314     0.96   0.338    -1.016685    2.957111
       _cons |   5.269006   7.600472     0.69   0.488    -9.648664    20.18668
------------------------------------------------------------------------------

      Source |       SS       df       MS              Number of obs =     872
-------------+------------------------------           F( 12,   859) =    2.09
       Model |  239.044831    12  19.9204026           Prob > F      =  0.0153
    Residual |  8172.22925   859  9.51365454           R-squared     =  0.0284
-------------+------------------------------           Adj R-squared =  0.0148
       Total |  8411.27408   871  9.65703109           Root MSE      =  3.0844
```

BPI_N02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21G17A	-.1565633	.0627212	-2.50	0.013	-.2796679	-.0334586
Q21G17B	.008049	.0616456	0.13	0.896	-.1129446	.1290426
Q21G17C	-.028924	.0606902	-0.48	0.634	-.1480425	.0901946
Q21G17D	-.1763435	.097224	-1.81	0.070	-.367168	.014481
Q21G17E	-.2436449	.0893163	-2.73	0.007	-.4189486	-.0683413
Q21G17F	.0723512	.0546897	1.32	0.186	-.0349899	.1796923
Q21G17G	.1412426	.0990663	1.43	0.154	-.0531977	.3356829
Q21G17H	-.0430551	.0571229	-0.75	0.451	-.1551718	.0690617
Q21G17I	-.0503049	.0591284	-0.85	0.395	-.1663579	.0657481
Q21G17J	-.2823957	.2961272	-0.95	0.341	-.8636132	.2988219
Q21G17K	-.0447917	.7738861	-0.06	0.954	-1.563721	1.474137
Q21G17O	.7052083	.7738861	0.91	0.362	-.8137208	2.224137
_cons	3.316692	5.810353	0.57	0.568	-8.087459	14.72084

The first thing I want to appreciate about the data is that, as one might expect, disciplinary methods from parents have the largest effect on a child's external behavior problems (r = .25), then on positive child behavior (r = .22), and then on internal behavior problems (r =.17). Now let's analyze the significant variables (using the 10% level) in each regression. I will label them as "good" or "bad" depending on whether each disciplinary technique had a positive or negative effect on child behavior. Here is the list in order of significance:
For POSBEH: ignoring (bad), grounding (bad), taking away allowance (good), sending to room (bad)
For External BPI: ignoring (bad), grounding (bad)
For Internal BPI: ignoring (bad), grounding (bad), giving chores (bad)

By far the most significant variables were the grounding and ignoring variables. Children whose parents incorporate these disciplinary techniques tend to be significantly worse-behaved than other children (p = 0.000 in both external behavior indices!). Another interesting thing to note is that parents who took their child's allowance away and those who incorporated "other" disciplinary methods were the only ones who significantly produced "good" results on all three measures, though only one of these was statistically significant. However, when we combine these three results, we can essentially claim statistical significance since the probability it would be positive at their levels on all three tests is below the 10% level. What's more is that these variables are essentially the only ones that are EVER positive. Overwhelmingly, discipline seems to be "bad" from a purely statistical point. Each variable tested, though rarely at a significant level, tended to have negative effects on child behavior. It seems strange to claim that all disciplinary methods are bad, but it seems that the majority of the ones inquired about seem to be ineffective. I also want to note that we can't claim any sort of confounding here because none of these questions asked how much a parent actually had to discipline their child, only what they would in the event of an outburst. Therefore, there is no bias in the sense that parents who discipline their children more are going to have worse-behaved children.

I guess these findings surprise me somewhat since grounding seems to be the most traditional method of discipline. The data here suggests that it is quite possibly a horrible method of discipline and that there needs to be a new (perhaps monetary) standard of discipline in American households, at least for

this age group. Unsurprisingly, parents who instituted alternative forms of discipline had better-behaved children, quite possibly because they've taken enough of an interest in their child's development to explore new avenues for disciplining rather than simply sticking to traditional methods. This could explain why grounding seems detrimental. It is possibly the easiest or most uninvolved form of punishment, and might reflect a parent who is not heavily invested in their child's well-being.

In addition to a parent's response to an outburst from their child, I wanted to see what ways parents could best deal with disappointment or poor performance from their child. There was a question asked to the parents of the 6-9 age group about how they'd respond if their child brought home subpar grades on a report card. My initial prediction is that, unlike the previous data, a parent's response to their child bringing home poor grades is going to explain internal BPI far more than external BPI, like the disciplinary methods one. My reasoning here is that a parent's response to their child's failure likely has much larger psychological implications on the child than it does behavioral (as their response to a temper tantrum might). Also, note that these variables (except for J) are scaled such that the higher they are, the more likely a parent is to institute this specific reaction to poor grades (unlike the previous measures we looked at).

The variables:

Q21G16A: Contact faculty	B: Talk with child	C: Watch activities more closely
E: Lecture child	F: Wait for improvement	G: Tell child to spend more time
H: Help child more	I: Limit activities	J: Other things

```
      Source |       SS       df       MS              Number of obs =     855
-------------+------------------------------           F(  9,   845) =    1.64
       Model |  4.60807732     9  .512008591           Prob > F      =  0.0986
    Residual |  263.136301   845  .311403906           R-squared     =  0.0172
-------------+------------------------------           Adj R-squared =  0.0067
       Total |  267.744378   854  .313518007           Root MSE      =  .55804

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21G16A |    .008186   .0173736     0.47   0.638    -.0259145    .0422866
     Q21G16B |   .0417549   .0522764     0.80   0.425     -.060852    .1443618
     Q21G16C |   .0376495   .0432502     0.87   0.384     -.047241      .12254
     Q21G16E |  -.0217681   .0147451    -1.48   0.140    -.0507094    .0071731
     Q21G16F |   .0133115   .0122679     1.09   0.278    -.0107676    .0373906
     Q21G16G |   -.022357   .0280079    -0.80   0.425    -.0773302    .0326162
     Q21G16H |   .0861925   .0459594     1.88   0.061    -.0040155    .1764006
     Q21G16I |  -.0009366   .0156339    -0.06   0.952    -.0316225    .0297493
     Q21G16J |   .0176203   .0110564     1.59   0.111    -.0040809    .0393214
       _cons |    3.39148    .266226    12.74   0.000     2.868939    3.914022
------------------------------------------------------------------------------
```

```
      Source |       SS       df       MS              Number of obs =     855
-------------+------------------------------           F(  9,   845) =    2.72
       Model |  408.251731     9  45.3613034           Prob > F      =  0.0039
    Residual |  14072.1296   845  16.6534078           R-squared     =  0.0282
```

```
-------------+------------------------------           Adj R-squared =  0.0178
       Total |  14480.3813   854    16.95595           Root MSE      =  4.0809

------------------------------------------------------------------------------
     BPI_E02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21G16A |  -.0767492   .1270516    -0.60   0.546    -.3261229    .1726245
     Q21G16B |  -.2893549   .3822922    -0.76   0.449    -1.039709    .4609988
     Q21G16C |   -.120643   .3162844    -0.38   0.703    -.7414383    .5001523
     Q21G16E |    .356309   .1078292     3.30   0.001     .1446645    .5679535
     Q21G16F |   .1596138   .0897138     1.78   0.076    -.0164743    .3357019
     Q21G16G |   .0119625   .2048188     0.06   0.953    -.3900508    .4139758
     Q21G16H |  -.5782318   .3360967    -1.72   0.086    -1.237914    .0814505
     Q21G16I |   .0618072   .1143293     0.54   0.589    -.1625955    .2862099
     Q21G16J |  -.0847995   .0808541    -1.05   0.295     -.243498    .0738989
       _cons |   9.380553   1.946883     4.82   0.000     5.559258    13.20185
------------------------------------------------------------------------------
```

```
      Source |       SS       df       MS              Number of obs =     855
-------------+------------------------------           F(  9,   845) =    1.18
       Model |   102.62349     9    11.40261           Prob > F      =  0.3036
    Residual |  8157.54727   845  9.65390209           R-squared     =  0.0124
-------------+------------------------------           Adj R-squared =  0.0019
       Total |  8260.17076   854   9.6723311           Root MSE      =  3.1071

------------------------------------------------------------------------------
     BPI_N02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21G16A |  -.0528668   .0967342    -0.55   0.585    -.2427342    .1370007
     Q21G16B |  -.0371004   .2910686    -0.13   0.899    -.6084025    .5342018
     Q21G16C |  -.0649439   .2408118    -0.27   0.787    -.5376033    .4077155
     Q21G16E |   .0635798   .0820987     0.77   0.439    -.0975615    .2247211
     Q21G16F |   .1061806   .0683061     1.55   0.120    -.0278889      .24025
     Q21G16G |   .0303605   .1559444     0.19   0.846    -.2757233    .3364442
     Q21G16H |  -.2900015   .2558964    -1.13   0.257    -.7922685    .2122656
     Q21G16I |   .0733177   .0870477     0.84   0.400    -.0975374    .2441728
     Q21G16J |  -.1134459   .0615605    -1.84   0.066    -.2342753    .0073834
       _cons |   4.812599   1.482313     3.25   0.001     1.903153    7.722046
------------------------------------------------------------------------------
```

I was fairly surprised at the lack of explanatory power in these measures, specifically in the internal BPI measures, but there are some significant findings here that I expected. Significant findings (in order of significance):

POSBEH: help child more (good), other things (bad), lecture child (bad)

External BPI: lecture child (bad), wait for improvement (bad), help child more (good)

Internal BPI: other things (bad), wait for improvement (bad)

The findings are fairly consistent and self-explanatory here. Like our previous findings, ignoring the issue (waiting for improvement) is associated with poor child behavior, probably because it indicates a parent who is not entirely invested in or interested in their child's life. However, even worse than this was lecturing a child (especially when it came to external BPI), which may display not only a distaste and

disappointment toward the child, but also an unwillingness to help the child along. Contrast this almost perfectly with the only "good" way (statistically) to respond to a child's poor grades: helping them more. Personally helping a child when they perform at a substandard level not only communicates an interest in the child's life, but it communicates to a child that they have a parent who is willing to make personal sacrifices in order to help them. This strategy stands in stark contrast and opposition to both lecturing and waiting for improvement, which communicate disinterest at best and disdain at worst. Though there were few statistically significant findings, it's worth noting that there is a distinguishable pattern among all three measures for most variables. For instance, note the following findings: contacting faculty, talking with the child, and watching their activities more closely produced consistently "good" results, while telling a child to spend more time on their work and limiting their activities produced consistently "bad" results, though none of these were at the statistically significant level.

<u>10+ Years of Age</u>

This age group will certainly produce the most interesting findings, not only because it's the largest group and therefore provides a wealth of data, but because parents employ the widest range of disciplinary techniques in this age category and a child's disciplinary issues are most visible. Here we will examine if the same disciplinary techniques that seemed most effective at the 6-9 age group still hold true for older children. Do different disciplinary techniques communicate different things to children of different ages, or are the same general methods consistent regardless of age? Here we look at both of the same categories of questions (parental response to an outburst as well as poor grades) as we did for 6-9 year olds and hope to discern what changes in disciplinary behavior are optimal for this age category. The variables:

Q21H17A: Ground child	B: Spank child	C: Talk with child
D: Give child chores	E: Ignore it	F: Send child to room
G: Take away allowance	H: Take away privileges	I: Give child "time out"
J: Hug or console child	K: Hit or threaten child	M: involve the police
N: Kick child out of house	P: Make child eat soap	Q: Give child counseling

```
  Source |       SS       df       MS              Number of obs =    1884
-------------+------------------------------           F( 15,  1868) =    2.45
       Model |  13.8940961    15  .926273073           Prob > F      =  0.0015
    Residual |  705.240935  1868  .377537974           R-squared     =  0.0193
-------------+------------------------------           Adj R-squared =  0.0114
       Total |  719.135031  1883  .381909204           Root MSE      =  .61444

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21H17A |   .0124825    .008056     1.55   0.121    -.0033171    .0282822
     Q21H17B |   .0128545   .0098729     1.30   0.193    -.0065086    .0322176
     Q21H17C |   .0023636   .0084669     0.28   0.780    -.0142421    .0189693
     Q21H17D |  -.0044004   .0111294    -0.40   0.693    -.0262278     .017427
     Q21H17E |   .0357805   .0123424     2.90   0.004     .0115742    .0599868
     Q21H17F |   .0184491   .0076691     2.41   0.016     .0034081    .0334901
     Q21H17G |  -.0052578   .0112379    -0.47   0.640    -.0272981    .0167825
```

Q21H17H	.009015	.0078319	1.15	0.250	-.0063453	.0243753
Q21H17I	.0149202	.0110435	1.35	0.177	-.0067386	.0365791
Q21H17J	.0585442	.069304	0.84	0.398	-.0773772	.1944656
Q21H17K	-.0481324	.0431313	-1.12	0.265	-.1327231	.0364583
Q21H17M	.084333	.0489148	1.72	0.085	-.0116004	.1802664
Q21H17N	.0245511	.054712	0.45	0.654	-.0827519	.1318541
Q21H17P	-.0221102	.1091247	-0.20	0.839	-.2361292	.1919089
Q21H17Q	.4040912	.1538716	2.63	0.009	.1023128	.7058696
_cons	1.228944	1.116251	1.10	0.271	-.9602853	3.418173

Source	SS	df	MS		Number of obs	= 1884
					F(15, 1868)	= 5.53
Model	1433.28176	15	95.5521176		Prob > F	= 0.0000
Residual	32302.9072	1868	17.2927769		R-squared	= 0.0425
					Adj R-squared	= 0.0348
Total	33736.189	1883	17.9161917		Root MSE	= 4.1585

BPI_E02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21H17A	-.060654	.0545218	-1.11	0.266	-.167584	.046276
Q21H17B	-.1608241	.0668187	-2.41	0.016	-.2918712	-.0297769
Q21H17C	.0720815	.0573032	1.26	0.209	-.0403035	.1844665
Q21H17D	-.1347315	.0753225	-1.79	0.074	-.2824566	.0129936
Q21H17E	-.1637467	.0835315	-1.96	0.050	-.3275716	.0000782
Q21H17F	-.2075154	.0519038	-4.00	0.000	-.3093108	-.1057199
Q21H17G	-.1207324	.076057	-1.59	0.113	-.269898	.0284332
Q21H17H	-.0299101	.0530056	-0.56	0.573	-.1338666	.0740463
Q21H17I	-.1356043	.0747408	-1.81	0.070	-.2821886	.0109801
Q21H17J	-.9583512	.4690404	-2.04	0.041	-1.878249	-.0384528
Q21H17K	.4958826	.2919074	1.70	0.090	-.0766163	1.068382
Q21H17M	-.7733532	.3310491	-2.34	0.020	-1.422618	-.1240882
Q21H17N	-.7660197	.3702835	-2.07	0.039	-1.492233	-.0398069
Q21H17P	-.5310178	.7385415	-0.72	0.472	-1.979471	.9174355
Q21H17Q	-3.004147	1.041383	-2.88	0.004	-5.046544	-.9617497
_cons	37.37663	7.554639	4.95	0.000	22.56021	52.19305

Source	SS	df	MS		Number of obs	= 1884
					F(15, 1868)	= 3.77
Model	574.857328	15	38.3238218		Prob > F	= 0.0000
Residual	18983.6904	1868	10.1625752		R-squared	= 0.0294
					Adj R-squared	= 0.0216
Total	19558.5478	1883	10.386908		Root MSE	= 3.1879

BPI_N02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21H17A	.0354875	.0417965	0.85	0.396	-.0464852	.1174602
Q21H17B	-.012202	.0512233	-0.24	0.812	-.1126629	.0882589
Q21H17C	-.0379221	.0439287	-0.86	0.388	-.1240766	.0482323
Q21H17D	.0203055	.0577423	0.35	0.725	-.0929408	.1335517
Q21H17E	-.1947749	.0640353	-3.04	0.002	-.3203632	-.0691865

```
    Q21H17F |   -.179115   .0397895    -4.50   0.000    -.2571515   -.1010784
    Q21H17G |  -.0334422   .0583054    -0.57   0.566    -.1477927    .0809083
    Q21H17H |  -.0113946   .0406342    -0.28   0.779    -.0910878    .0682985
    Q21H17I |  -.0316665   .0572964    -0.55   0.581    -.1440383    .0807052
    Q21H17J |  -.8432231    .359567    -2.35   0.019    -1.548418   -.1380279
    Q21H17K |   .2227088   .2237766     1.00   0.320    -.2161696    .6615872
    Q21H17M |  -.1935917   .2537826    -0.76   0.446     -.691319    .3041357
    Q21H17N |  -.5112733   .2838598    -1.80   0.072    -1.067989    .0454424
    Q21H17P |  -.1991408   .5661669    -0.35   0.725    -1.309527    .9112453
    Q21H17Q |   -2.82223   .7983258    -3.54   0.000    -4.387935   -1.256526
      _cons |   26.66845   5.791396     4.60   0.000     15.31017    38.02674
------------------------------------------------------------------------------
```

Interestingly, the POSBEH variable has moved from being very well described by disciplinary measures at the 6-9 age group level to being the least described variable at 10+ years of age. However, the BPI variables are still explained pretty well. It is worth noting that every single measure, though we are including significantly more variables, is less described than it was with fewer variables at the 6-9 age level. This is to be expected, as children's behavior becomes influenced by a great number of sources besides their parents' discipline as they become older. Children become more and more influenced by their friends' behavior as well as developing their own notions or morality and good behavior beyond what their parents taught them due to more life experience. It therefore makes sense that even though we're including more variables, we lose considerable explanatory power with this age group. Since we are including 15 explanatory variables for this age group, I will move my significance level down to 5%.

The significant findings (in order of significance):
POSBEH: ignore it (bad), counseling (bad), send child to room (bad)
External BPI: send child to room (bad), counseling (bad), spanking (bad), involving police (bad), kicking out of house (bad), hug/console (bad), ignore it (bad)
Internal BPI: send child to room (bad), counseling (bad), ignore it (bad), hug/console (bad)

These findings are pretty depressing, to be honest. There was not a single positive disciplinary response to be found among this age group. All of these disciplinary techniques seem to result only in increased misbehavior among children. When contrasting these findings with the 6-9 age range, it's interesting to see that though grounding still seems to be associated with mostly negative behaviors, it is not nearly as strong as it was at the 6-9 age group level and never achieves statistical significance here. Clearly it is not as detrimental of a punishment at this older age group. I still have no idea why sending a child to their room is essentially the worst fate a parent can give as far as predicting their child's behavior problems, but it is without a doubt the most statistically significant finding for BPI measures, and seems to just be an absolutely horrendous thing to do. It also appears that sending a child to counseling has quite poor results as well. This could be because, like sending a child to their room, it involves very little personal parental involvement or acknowledgement (much like ignoring or involving the police or hugging them). The general trend seems to be that, the more active parent involvement there is required in a given disciplinary action, the better it is. The worst variables tend to be the ones that require the least effort or interest on the part of the parent. There is, however, one thing I'm very

disturbed about with these findings. The only variable that consistently produced "good" results was the one where parents said they would restrain, hit, or threaten their child for such an outburst. This showed positive (but never significant) results on all three measures, which seems quite odd to me, but I cannot ignore the data. I offer no great explanation.

I then moved on to see what sorts of behaviors would be the best response to a child's poor grades on a report card in this age group. My initial assumption is that talking to the child would be best, whereas lecturing will still be the worst thing (as it was at the 6-9 age level). The variables:

Q21H16A: Contact faculty	B: Talk with the child	C: Watch activities
D: Punish child	E: Lecture child	F: Wait for improvement
G: Tell child to spend more time	H: Help child more	I: Limit activities
J: Other things		

```
      Source |       SS       df       MS              Number of obs =    1793
-------------+------------------------------           F( 10,  1782) =    6.79
       Model |  24.1483916    10  2.41483916           Prob > F      =  0.0000
    Residual |  634.042631  1782  .355803946           R-squared     =  0.0367
-------------+------------------------------           Adj R-squared =  0.0313
       Total |  658.191023  1792  .367294097           Root MSE      =  .59649

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21H16A |   .0147277   .0111608     1.32   0.187     -.007162    .0366174
     Q21H16B |   .1747949   .0383849     4.55   0.000     .0995108     .250079
     Q21H16C |   .0057291   .0270428     0.21   0.832    -.0473097     .058768
     Q21H16D |  -.0272701   .0108964    -2.50   0.012    -.0486412   -.0058991
     Q21H16E |  -.0415516   .0129218    -3.22   0.001     -.066895   -.0162081
     Q21H16F |   .0322743   .0089913     3.59   0.000     .0146398    .0499089
     Q21H16G |   -.024759   .0219997    -1.13   0.261    -.0679069    .0183888
     Q21H16H |   .0138219    .015699     0.88   0.379    -.0169685    .0446122
     Q21H16I |   .0360065   .0120447     2.99   0.003     .0123833    .0596297
     Q21H16J |   .0087326    .007737     1.13   0.259    -.0064421    .0239072
       _cons |   3.222981   .1940033    16.61   0.000     2.842483    3.603479
------------------------------------------------------------------------------

      Source |       SS       df       MS              Number of obs =    1793
-------------+------------------------------           F( 10,  1782) =    6.07
       Model |  1054.61552    10  105.461552           Prob > F      =  0.0000
    Residual |  30979.3354  1782  17.3845878           R-squared     =  0.0329
-------------+------------------------------           Adj R-squared =  0.0275
       Total |  32033.9509  1792  17.8760887           Root MSE      =  4.1695

------------------------------------------------------------------------------
     BPI_E02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21H16A |  -.0176848   .0780143    -0.23   0.821    -.1706939    .1353243
     Q21H16B |  -.3451593     .26831    -1.29   0.198    -.8713946     .181076
     Q21H16C |  -.4243179   .1890287    -2.24   0.025    -.7950592   -.0535766
```

```
     Q21H16D |   .3141165   .0761657     4.12   0.000      .164733    .4634999
     Q21H16E |   .3107684   .0903232     3.44   0.001     .1336178     .487919
     Q21H16F |  -.0351467    .062849    -0.56   0.576    -.1584122    .0881189
     Q21H16G |   .0806396   .1537776     0.52   0.600    -.2209637    .3822429
     Q21H16H |   .0706509   .1097359     0.64   0.520    -.1445737    .2858754
     Q21H16I |  -.1131029   .0841923    -1.34   0.179    -.2782289    .0520232
     Q21H16J |  -.0789423   .0540818    -1.46   0.145    -.1850128    .0271281
       _cons |   7.295797   1.356082     5.38   0.000     4.636118    9.955475
------------------------------------------------------------------------------

      Source |       SS       df       MS              Number of obs =    1793
-------------+------------------------------           F( 10,  1782) =    1.39
       Model |  141.192244    10  14.1192244           Prob > F      =  0.1786
    Residual |  18098.0727  1782  10.1560453           R-squared     =  0.0077
-------------+------------------------------           Adj R-squared =  0.0022
       Total |  18239.2649  1792  10.1781612           Root MSE      =  3.1869

------------------------------------------------------------------------------
     BPI_N02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21H16A |   .0844338   .0596285     1.42   0.157    -.0325154     .201383
     Q21H16B |  -.2690572    .205077    -1.31   0.190    -.6712739    .1331595
     Q21H16C |  -.2228874   .1444801    -1.54   0.123    -.5062556    .0604808
     Q21H16D |   .0146667   .0582156     0.25   0.801    -.0995113    .1288448
     Q21H16E |   .0620739   .0690366     0.90   0.369    -.0733274    .1974752
     Q21H16F |  -.0176038   .0480373    -0.37   0.714    -.1118192    .0766116
     Q21H16G |   .0187336   .1175366     0.16   0.873    -.2117904    .2492577
     Q21H16H |   .1742852   .0838743     2.08   0.038      .009783    .3387875
     Q21H16I |  -.0562881   .0643506    -0.87   0.382    -.1824987    .0699225
     Q21H16J |  -.0445294   .0413363    -1.08   0.282    -.1256021    .0365433
       _cons |   4.456041   1.036492     4.30   0.000     2.423173    6.488909
------------------------------------------------------------------------------
```

The most glaring result regarding descriptive power here is two-fold: first that the explanatory power of these regressions are much stronger than they were at the 6-9 age group. At the 6-9 age level, how a parent responds to poor grades was hardly any sort of indicator of child behavior, but now it indicates quite a bit and we have many more statistically significant findings. Secondly, and quite surprising to me, was that a parent's response to grades seems to play itself out way more in external behavior problems than internal. This is surprising to me since I would usually assume that there would more psychological damage due to a parent's disappointment rather than outward behavioral issues. Significant findings (in order of significance):
POSBEH: talk with the child (good), wait for improvement (good), lecture child (bad), limit activities (good), punish child (bad)
External BPI: punish child (bad), lecture child (bad), watch activities (good)
Internal BPI: help child more (bad)

It's really interesting to see how waiting for improvement went from being bad at the 6-9 age group to unanimously positive at the 10+ age group. As expected, punishing and lecturing a child were the worst things a parent could do in response to poor grades, while talking with the child about it was the best thing a parent could do. Once again, this data reinforces the major theme that the best response to the child is what most indicates an active interest in the child's life. This is most obviously seen in the

contrast between talking to the child (the best thing you could do) and lecturing the child (the worst thing you could do). They both involve talking, but the former indicates to the child that their opinion and perspective truly matters while the latter implies that the parent is really only concerned in how their child's performance reflects on their own parenting ability rather than take a keen interest in the child's personal struggles. Also, while it appears that waiting for improvement may have indicated a disinterest in a parent in the 6-9 age group, in this 10+ age group, it appears to be a very positive thing, quite likely communicating a patience and understanding that the child is having difficulty rather than a complete lack of care. On the flip side, giving a child more help does not help nearly as much as it did at the 6-9 age group, and is associated with significant increases in sad or withdrawn behavior in children at the 10+ age group. Clearly, helping at this age does not communicate as much of an interest in assisting as much as it does that a child is not capable of achieving on their own. This is of course the entirely opposite effect of what happened at the 6-9 age level.

In general, we've seen that there is quite a large paradigm shift that occurs in psychological effects of different parenting styles at different ages. There are a couple rules of thumb that can be applied throughout, such as "don't lecture your child when they get bad grades", "when your child acts up, do not ignore them", and "sending your child to their room generally doesn't help much", but many other rules seem only to apply at a given age range. For example, grounding a child who is between 6 and 9 years of age seems to be an entirely inappropriate disciplinary technique as it is related to child behavior, but at the 10+ age level, it is not nearly as detrimental. In addition, helping a child with their work seems to be a great response to low performance at the younger age group, while this seems entirely inappropriate at the older age range. Also, it seems that talking to your 6-9 year-old about their poor report card, though somewhat helpful, doesn't make much of a difference, while doing this with your 10+ year-old seems to be the best thing you can possibly do.

Positive Reinforcement

In addition to disciplinary methods, I wanted to see if I could discern which sorts of positive parental involvement was most influential on a child's psychological and behavioral well-being.

<u>6-9 Years of Age</u>

The variables:

Q21G15D: number of times parent praised child in the last week

Q21G15F: number of times parent showed child physical affection in the last week

Q21G21A1: child's allowance in dollars

ER24099: total family income

Inc2: total family income squared

```
      Source |       SS       df       MS              Number of obs =     386
-------------+------------------------------           F(  5,   380) =    3.46
       Model |  5.47919906     5  1.09583981           Prob > F      =  0.0045
    Residual |  120.217102   380  .316360795           R-squared     =  0.0436
-------------+------------------------------           Adj R-squared =  0.0310
       Total |  125.696301   385  .326483899           Root MSE      =  .56246
```

POSBEH02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21G15D	.0253311	.0074983	3.38	0.001	.0105876	.0400745
Q21G15F	-.0012269	.0035961	-0.34	0.733	-.0082976	.0058438
Q21G21A1	-.0041917	.0034869	-1.20	0.230	-.0110478	.0026644
ER24099	6.32e-07	9.39e-07	0.67	0.502	-1.22e-06	2.48e-06
inc2	-1.50e-13	1.78e-12	-0.08	0.933	-3.66e-12	3.36e-12
_cons	4.066205	.067483	60.26	0.000	3.933518	4.198892

BPI_E02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21G15D	-.1387468	.0541592	-2.56	0.011	-.245236	-.0322575
Q21G15F	-.0008676	.0259739	-0.03	0.973	-.0519382	.0502029
Q21G21A1	.0112035	.0251855	0.44	0.657	-.0383169	.0607238
ER24099	-.0000267	6.79e-06	-3.93	0.000	-.00004	-.0000133
inc2	3.51e-11	1.29e-11	2.72	0.007	9.73e-12	6.04e-11
_cons	7.801077	.4874173	16.00	0.000	6.842704	8.759449

BPI_N02	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Q21G15D	-.083978	.0408828	-2.05	0.041	-.164363	-.0035931
Q21G15F	-.0083962	.0196068	-0.43	0.669	-.0469475	.0301551
Q21G21A1	.0004996	.0190116	0.03	0.979	-.0368815	.0378807
ER24099	-.0000148	5.12e-06	-2.88	0.004	-.0000248	-4.69e-06
inc2	1.38e-11	9.73e-12	1.42	0.157	-5.34e-12	3.29e-11
_cons	4.425436	.3679338	12.03	0.000	3.701995	5.148877

I admit that I was fairly surprised by these findings. It seems that once we control for how much a parent praises a child 6-9 years of age, physical affection and allowance amount seem to have no impact whatsoever. I was also quite astounded at just how crucial praising a child is in each of these measures. It predicted more positive behavior and less internal and external behavioral problems at the 2% level in each regression. It's also interesting to note that in the positive behavior regression (where results seem to be magnified for some reason), higher allowance had a pretty convincing negative effect on a child's positive behavior. I also found it very interesting that family income seems to play little to no role in a child's helpful behaviors, but plays an outstandingly significant role in the presence of behavioral problems. Clearly children in lower income households tend to exhibit significantly more behavior problems, though their outward positive behavior is practically indistinguishable from other children. Also interesting was that behavior problems increase with income squared, indicating that behavior problems are a decreasing but convex function of income, meaning that at a certain point, more income stops becoming important or even becomes detrimental regarding behavioral issues. However, a parent's praise is without a doubt the best predictor of positive behavior at this age.

<u>10+ Years of Age</u>

My assumption here is that allowance is going to become a bit more important at this age because it communicates more to a child how much they are trusted or valued than it would at the 6-9 age range, especially because children are now at the point where they'll be comparing their allowance

or income to that of other students in order to judge how good or fair their parents are, and maybe even how much they love or appreciate them. I also expect a decrease in the importance of praising a child. The variables:
Q21H15D: number of times parent praised child in the last week
Q21H15F: number of times parent showed child physical affection in the last week
Q21H15H: number of times parent spoke positively about child to someone else in the last week
Q21H22B: child's allowance in dollars
ER24099: total family income
Inc2: total family income squared

```
    Source |       SS       df       MS              Number of obs =     814
-------------+------------------------------           F(  6,   807) =    1.98
       Model |  4.35342489     6  .725570815           Prob > F      =  0.0667
    Residual |  296.460198   807  .367360841           R-squared     =  0.0145
-------------+------------------------------           Adj R-squared =  0.0071
       Total |  300.813623   813  .370004457           Root MSE      =   .6061

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21H15D |   .0109739   .0072162     1.52   0.129    -.0031909    .0251386
     Q21H15F |  -.0077234   .0036475    -2.12   0.035    -.0148831   -.0005637
     Q21H15H |   .0129297   .0065191     1.98   0.048     .0001334    .0257261
     Q21H22B |  -.0005862   .0006358    -0.92   0.357    -.0018342    .0006618
     ER24099 |  -1.36e-07   4.81e-07    -0.28   0.777    -1.08e-06    8.09e-07
        inc2 |   1.26e-14   2.78e-13     0.05   0.964    -5.33e-13    5.59e-13
       _cons |   4.148548   .0451401    91.90   0.000     4.059942    4.237154
------------------------------------------------------------------------------
------------------------------------------------------------------------------
     BPI_E02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21H15D |  -.0658602   .0502745    -1.31   0.191    -.1645445     .032824
     Q21H15F |   .0397433   .0254116     1.56   0.118    -.0101373    .0896239
     Q21H15H |  -.0601975   .0454177    -1.33   0.185    -.1493482    .0289532
     Q21H22B |   .0032793   .0044295     0.74   0.459    -.0054155    .0119741
     ER24099 |  -.0000138   3.35e-06    -4.13   0.000    -.0000204   -7.26e-06
        inc2 |   5.35e-12   1.94e-12     2.76   0.006     1.54e-12    9.15e-12
       _cons |   6.489582   .3144858    20.64   0.000     5.872276    7.106889
------------------------------------------------------------------------------
------------------------------------------------------------------------------
     BPI_N02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Q21H15D |   .0348166   .0378702     0.92   0.358     -.039519    .1091523
     Q21H15F |   .0298181   .0191417     1.56   0.120    -.0077554    .0673916
     Q21H15H |  -.0724163   .0342117    -2.12   0.035    -.1395707   -.0052619
     Q21H22B |    .004388   .0033366     1.32   0.189    -.0021615    .0109375
     ER24099 |  -4.31e-06   2.53e-06    -1.71   0.088    -9.27e-06    6.49e-07
        inc2 |   1.44e-12   1.46e-12     0.99   0.325    -1.43e-12    4.30e-12
       _cons |   3.176728   .2368921    13.41   0.000      2.71173    3.641725
------------------------------------------------------------------------------
```

Here we see the same general effects for income as we did before (significant decreasing convex effect on BPI and essentially no effect on POSBEH), while we see a very different story in our other variables. In

this case, praising a child has a notable but not statistically significant positive effect on positive child behavior and external behavior problems, but has a negative effect on internal behavioral problems. Now we are starting to see some of the strange tendencies discussed at the beginning of this paper (it seems we've found our culprit age group in addition to the specific parent-child behavioral culprits). Even more interestingly, parental physical affection at this age group proved to be, completely without fail, statistically significantly detrimental in every measure. This is probably one of the most surprising findings of my paper. For whatever reason, the more affection parents are toward their 10+ year-olds, the more likely these children are to act out and demonstrate sad or withdrawn tendencies. Also, I could see how we would be at a lack for knowledge regarding the direction of causation for internal BPI scores (perhaps parents hug their kids more because they're always sad), but in the other two external measures, there seems absolutely no reason why parents would be more affectionate toward children who display delinquent tendencies. For this reason, I'm quite convinced that the causation arrow most likely points the other direction: increased physical affection at this age likely causes children to act out in order to assert their independence, which is called into question when parents show the type of affection they would show to younger children.

The last interesting finding is that how much a parents praise their child to others seems to be a greater predictor of their child's behavior than how much they praise the child to their face (this is astoundingly true in the internal BPI scores). I see the reason for this being that parents telling others about their children is a great indicator of how truly proud they are of their kids, which has a tremendous effect on internal BPI factors. My hypothesis is that children can tell when their parents are truly proud of them and when they are giving them mere lip service. As a result, giving your child words of affirmation may not help much if you're not telling anyone else about them, because how proud you truly are of your child is probably coming out in other ways within the house, and a child is not oblivious to this.

Parenting by Example

Giving

Lastly I wanted to study the effects of positive parental behavior outside the parent-child dynamic on a child's development. To do this, I wanted to study parents' behavior in philanthropic areas, first in giving, and then in service, in order to track how much a parent's example influences what kind of citizen their child becomes. The main problem I had with the data is that a value of 0 was assigned both to people who were not eligible for the question as well as those who didn't donate anything. Therefore, I had to drop the "0" values even though some of them would have been valuable (the people who did not give but were eligible). In addition, I could therefore only use two types of donations or else I would have dropped every data observation. The variables:
ER2483: Religious donations in dollars
ER23495: Donations to the needy in dollars

```
      Source |       SS       df       MS              Number of obs =     522
-------------+------------------------------           F(  4,   517) =    5.26
       Model |  5.08830714     4  1.27207678           Prob > F      =  0.0004
    Residual |  125.023687   517  .241825314           R-squared     =  0.0391
```

```
-------------+------------------------------           Adj R-squared =  0.0317
       Total |  130.111995   521  .249735114           Root MSE      =  .49176

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     ER23483 |  -6.05e-06   6.52e-06    -0.93   0.354    -.0000189    6.77e-06
     ER23495 |    .000028   .0000308     0.91   0.364    -.0000325    .0000885
     ER24099 |  -1.88e-07   1.45e-07    -1.29   0.196    -4.73e-07    9.72e-08
    WARMTH02 |   .1513466   .0366794     4.13   0.000     .0792876    .2234055
       _cons |    3.58642   .1491184    24.05   0.000     3.293467    3.879372
```

Clearly a parent's example (at least in giving) is not nearly as strong a predictor as their personal interactions with a child in producing people who contribute to society. Interestingly enough, religious donations even went the negative direction on child behavior (perhaps due to some of the trends we saw earlier regarding children who didn't want to go to church but were still forced). It seems that any parent can write a check. Children aren't paying nearly as close attention to that as they are to a parent's interactions with them personally.

Service

I encountered the same obstacles with service variables. A zero was both a person who didn't volunteer and someone who was ineligible to answer the question, thereby robbing me of tons of precious data if I threw out the observations, or completely skewing the data if I left them in. I chose the former to be narrow rather than inaccurate, but hoping to still glean some valuable information. The variables:

ER23628: Number of times head of household volunteered with the needy during the year

```
      Source |       SS       df       MS              Number of obs =      95
-------------+------------------------------           F(  3,    91) =    2.19
       Model |  1.79392668     3  .597975561           Prob > F      =  0.0943
    Residual |   24.823548    91  .272786242           R-squared     =  0.0674
-------------+------------------------------           Adj R-squared =  0.0367
       Total |  26.6174747    94  .283164624           Root MSE      =  .52229

------------------------------------------------------------------------------
    POSBEH02 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     ER23628 |  -.0053561    .006569    -0.82   0.417    -.0184045    .0076924
     ER24099 |  -1.51e-06   6.81e-07    -2.22   0.029    -2.86e-06   -1.60e-07
    WARMTH02 |   .1274995   .0887394     1.44   0.154    -.0487705    .3037694
       _cons |   3.808436   .3599744    10.58   0.000     3.093391    4.523481
```

Obviously with such few data points it's hard to really draw any conclusions (apparently only 95 people volunteered more than once). It does seem, however, that there is really no relationship at all between a child's positive behavior and the amount their parent volunteers once we have established the fact that their parent is in fact someone who already volunteers. It's interesting that there's even a slightly negative relationship, suggesting that maybe the parents should stay home more if they've given too much time to volunteer work. Most of the findings in this section were largely inconclusive, largely due

to the nature of the data, but also due to the fact that it just seems that a parent's example is not as influential as a parent's personal interactions with a child on that child's development. With this established, we've now come to a more holistic view on the needs of children from their parents and what things seem to be most important and what things just don't really seem to be that necessary according to the data.

Conclusion

After looking at the most apparent aspects of the parent-child relationship, there seems to be a fairly large set of common themes coming through from the data, some of which I could have easily predicted, but many more of which I was quite surprised by. Perhaps my greatest surprise in this entire project was how many times the data caught me completely off guard. I will now list what I consider to be the major findings of this research endeavor.

Finding #1: Predicting external behavior is far easier than predicting internal behavior

It seemed that every time I regressed a set of variables separately against positive behavior, external behavior problems, and internal behavior problems, the internal behavior problems always had the lowest R-squared values. This variable just seems to be far more complicated than the others and is simply less predictable from common economic variables. This is not an entirely intuitive fact, as it seems that external behavior is an equally difficult thing to predict and is usually thought to be a mere reflection of what lies beneath. Many instances where I expected a parent's interaction variables to affect mostly a child's internal well-being actually turned out describing a child's outward behavior much more accurately.

Finding #2: Physical affection, words of endearment, and relational involvement between parent and child are not all they're cracked up to be

Statistically speaking, parents talking with their children about relationships, telling them they're loved, and displays of physical affection do not predict any sort of benefit in child development, and often predict a wealth of internal psychological strife. It started with my strange finding of the opposite relationship between parental warmth and the internal BPI scores when controlling for externals, and then found its resolution in my study of physical affection being of no explanatory power at the 6-9 age range and seemingly quite detrimental at ages 10 and above. Though most parents and children alike would agree that physical affection is indeed a wonderful thing, the data shows that it is not nearly as important to child as other forms of attention that their parent might give to them.

Finding #3: Learning about and participating in a child's interests and activities is crucial to their development

The data showed that parents who take active roles in their children's lives, specifically in the area of play, have children who are significantly more well-adjusted, both in terms of displaying helpful and constructive external behavior as well as being psychologically healthy. This phenomenon also showed itself in the area of discussing relevant news topics of interest to the child. The main lesson to be learned here is that a child whose parents communicate that what is important to the child is also

important to them will be significantly more well-adjusted than his/her peers. In this way, some of the best investments a parent can make are in involving themselves in what is most exciting to the child so that they become aware that their parents are truly on their team in life. A child needs to know not only that their parents love them, but that they like them and are interested in their life. In fact, the data seems to demonstrate quite clearly that if a child feels liked by their parents, they will feel loved, but many children can feel loved without feeling liked, and this fact will play itself out negatively in their behavior.

Finding #4: Children are much more than a product of parenting strategies

This fact came through in multiple ways throughout the study. Perhaps the most obvious way is that I never once saw an R-squared value above 0.1, even when including over 15 explanatory variables. This only goes to say that even when we include every conceivable survey-identifiable parenting strategy into the mix, we cannot explain anything beyond 10% of a child's outward development. This goes to show that a child's development truly is a complex mix of factors. Though we have proven statistically that there are significant effects of parenting strategies on how children develop socially, the sum total of these verified effects has never amounted to anything more than 10% of what a child actually does. In addition to this, we saw that a child's personal religious convictions plays a tremendous role in their outward behavior, and this fact can only be slightly explained by their upbringing. Children, like the rest of human beings, do have a personal moral compass that is formed by numerous forces far outside the household.

Finding #5: Disciplinary methods must evolve over time

This one seems fairly obvious, and I don't know of any parents whose parenting strategies don't actually evolve over time, but what was most surprising in the data is that a similar disciplinary method, evoked at two different times in a child's life, can have opposite effects. It's natural to believe that a certain disciplinary method can evolve from optimal to less-than-optimal or vice versa, but I found that the swing can be as drastic as optimal to detrimental and vice versa.

Finding #6: Most traditional disciplinary methods seem to have negative effects on child behavior

I imagine that this must be one of the most controversial of my findings, and I'm not sure that I myself know quite what to do with it. But I cannot deny the data; it is clear that when we examine the sum total of different disciplinary strategies inquired about on the PSID, just about every single one of them indicated a more poorly-behaved child if a parent chose to incorporate that certain disciplinary strategy. What makes this so much more convincing is that the question asked was entirely hypothetical, meaning that there was no confounding due to parents needing to invoke more discipline due to poorly-behaved children. Interestingly, a child whose parent answered "no" to every question about disciplinary methods would have probably had the highest predicted score in terms of positive behaviors and low behavioral problems. It may just be that the traditional methods are not accomplishing what they were fashioned to do.

Finding #7: Ignoring a child who is throwing a temper tantrum is a horrible idea

The one universally horrible parenting strategy seemed to be the "ignore it" response when asked what a parent would do if their child was extremely upset. No matter what the age category, parents who invoked this disciplinary method tended to have the worst behaved children of all.

Finding #8: Sending children to their room (especially those who are very young or very old) is a really bad idea

For whatever reason, this strategy was even worse than ignoring a child for children below the age of six and above the age of ten. It seems to be a bad strategy because it implies a sense of carelessness on the part of the parent. For one, it doesn't really require any disciplinary effort on the part of the parent, secondly it may not be a punishment at all if the child really enjoys their bedroom, and thirdly (and most importantly), a child might feel that their parent is clueless, oblivious, and careless if they think that sending them to their room would be a good disciplinary method, and this message will likely cause a child to lash out to draw the genuine attention of their parent.

Finding #9: Grounding is not an appropriate disciplinary method for children ages 6-9

Parents who grounded their 6-9 year-olds tended to have really poorly-behaved kids. It could be that only the really badly-behaved kids did things horrible enough to warrant a parent's grounding, but more likely this sort of disciplinary method does absolutely nothing in order to correct a child's behavioral issues at this age.

Finding #10: Praising a child is incredibly important at the 6-9 age range

The amount a parent praised their 6-9 year-old completely demolished other variables such as physical affection or family income in predicting child development. This fact almost completely disappeared at older age ranges, but at this beginning stage in development, children seem to desperately need verbal affirmation in order to grow in the sort of positive identity they need to flourish psychologically.

Finding #11: A parent's example as a constructive citizen will have little to no effect on child development without helpful parenting methods to support

Another phrasing of this is that a parent's actions toward their child communicates infinitely more to a child, or has a significantly more profound influence on their development, than their actions within society as a whole. Model behaviors such as community service or giving to charity had absolutely no relationship to child development once we controlled for parental warmth variables, or those which identified parent-to-child interaction specifically. Therefore, most parents would be foolish to say that they are being a good parent to their child simply by showing them by example how they are to function in society. This sort of behavior, outside of meaningful one-on-one interaction between parent and child seems to be almost useless in predicting child development.

Finding #12: The greatest single predictor of child development is having parents who take a genuine interest in their child's life

This is probably the single finding that best summarizes this research. This came through in every aspect of the paper. Lecturing a child who brought home poor grades tended to only make issues

worse, while parents who talked to their children about them tended to have kids who were much more well-adjusted. Children whose parents ignore problems tended to have the worst behavioral scores of all. The key issue to be taken from many of the differences in findings regarding disciplinary methods was that the more a given method showed that a parent truly loved and cared for their child and had an interest in their problems and helping them rather than just reprimanding or "fixing" them, the more likely a child was displaying healthy behaviors. The more "hands-off" disciplinary approaches like sending a child to counseling or sending them to their room tended to predict poor behavior. In stark contrast, the hands-on approaches like helping a young child (6-9) with their school work when they bring home poor grades are the best indicators of positive development. All other findings in the paper are best understood if they're viewed through the lens of this realization that active parenting (though not overactive) is truly the best kind there is.

The Determinants of Consumer Sentiment in the Housing Market

Russell Bittmann

Mike Filicicchia

Duke Schaeffer

Introduction

The motivation for this research was to unearth some of the insights that might be gained from the consumer confidence research data in circulation, specifically regarding the housing market. Given the recent volatility and perceived importance of the housing market, we wanted to examine this confidence data to see what additional insight about market dynamics can be gained from this data set.

We primarily used the University of Michigan's Survey of Consumers to conduct our research. Our immediate purpose was not so much to use this data in order to predict real market trends or assess the accuracy with which consumers assess the market, but rather to unearth the determinants of consumer sentiment. In other words, what factors truly drive how consumers think of the markets, and how many of these factors are real or important, and how many are fictional or unimportant?

Our paper is divided in four sections. Section I introduces the structure of the Survey of Consumers and presents summary statistics of the relevant data. We introduce the basic data sets we use throughout the paper and indicate some of the operations we performed on the data to make it easier to analyze. Our discussion focuses on the general optimism in both buying and selling sentiment throughout the sample period, as this information is critical to understanding the rest of the findings throughout the paper. We postualte driving factors for the observed co-movement of buying and selling sentiment and break down the components of this co-movement. Also, in order to understand periods where summary buying and selling statistics were a little less defined, we look at some determinants for uncertainty in the housing market in order that we might better understand the data ascertained from the periods of relative sentimental certainty throughout the rest of the paper.

Due to similar patterns we found in buying and selling data, in Section II we compared the summary statistics further by looking at the reason response variables for each. These reason were the categorized responses consumers gave when asked why they believed it to be a good or bad time to buy or sell. After some analysis, we isolate interest rates and house prices as the two main determinants of consumer sentiment in the housing market, respectively. With this in mind, our group sought to first create variables that represented the consumer perceptions of these two quantities as they are translated into actual buying and selling sentiment. Once we created these variables, it became a priority to examine what market factors drove these specific perceptions, and specifically how closely they were related to the actual level of house prices and interest rates. When we discovered that this could not be undergone successfully throughout the whole period we studied (1992-2008), we looked for structural breaks in how consumer sentiment related to actual market fundamentals. After finding these important breaks, we generalized findings about how closely consumer sentiment follows market fundamentals and specifically looked into how well consumers account for inflationary effects when evaluating prices and interest rates.

In section III we attempt to explain the structural breakage found in Section II with underlying market factors and create a model that predicts consumer sentiment in the housing market over the entire fifteen-year scope of our data, clearing up all the muddying effects of earlier structural breaks.

In Section IV, we attempt to explain the inevitably of the recent housing crash due to the findings in this paper. Specifically, we focus on the disparity in price estimates people held when asked about buying and selling. When asked about buying, consumers disproportionately answered that prices were low, while when asked about selling, they tended to answer that prices were high (meaning that it was generally a good time to both buy and sell because of prices). We predicted that this unrestrained optimism, largely driven by deceptive mortgage terms, drove the price of housing upward as consumers continued to think that expensive houses were affordable, while sellers, though content with current price levels, upped their ask prices to accommodate the increased demand. This disparity continued at a very consistent level until the crash, when the trend reversed. We also look at the disparities in sentiment between homeowners and nonhomeowners.

I. "The Boom" – Consumer Sentiment Indexes and the Comovement of Buying and Selling Attitudes

The Survey

Buying and Selling Conditions Indexes

The most compelling component of the Survey of Consumers was the buying and selling conditions indexes. Each index was computed based on the response to the following set of questions (insert the appropriate buy/sell wording for the associated index):

(1) Generally speaking, do you think now is a good time or a bad time to buy/sell a house?

(2) Why do you say so? (Are there any other reasons)?

The second question was given in two parts where respondents would first answer the "why" question, and then later have to provide any additional reasons they could think of. Since respondents were allowed to cite multiple underlying reasons to their first response, the percentage of people in each group citing each reason would often add up to over 100%. Based on responses to this second question, the survey divided the reasons into 10 generalized categories or determinants (though an eleventh was added in November 1992).

The final buying and selling indexes are computed based on the response to the first question according to the following formula:

Buying/Selling Condition Index = (% good) – (% bad) + 100.

We focus on these indexes in the first section of this paper. Although this general index gives some insight into consumer sentiment, we were really interested in what drives this number. Thus, our research was considerably more focused on the second question, which is discussed in the second section of this paper.

Non-Response

Keep in mind that the index measure does not include in any way those who were uncertain of buying or selling conditions. Although data is collected on this subgroup, this sort of response is in no way counted toward the buying and selling conditions index. We do analysis on this subgroup later.

Time Scope of Buying and Selling Conditions Indexes

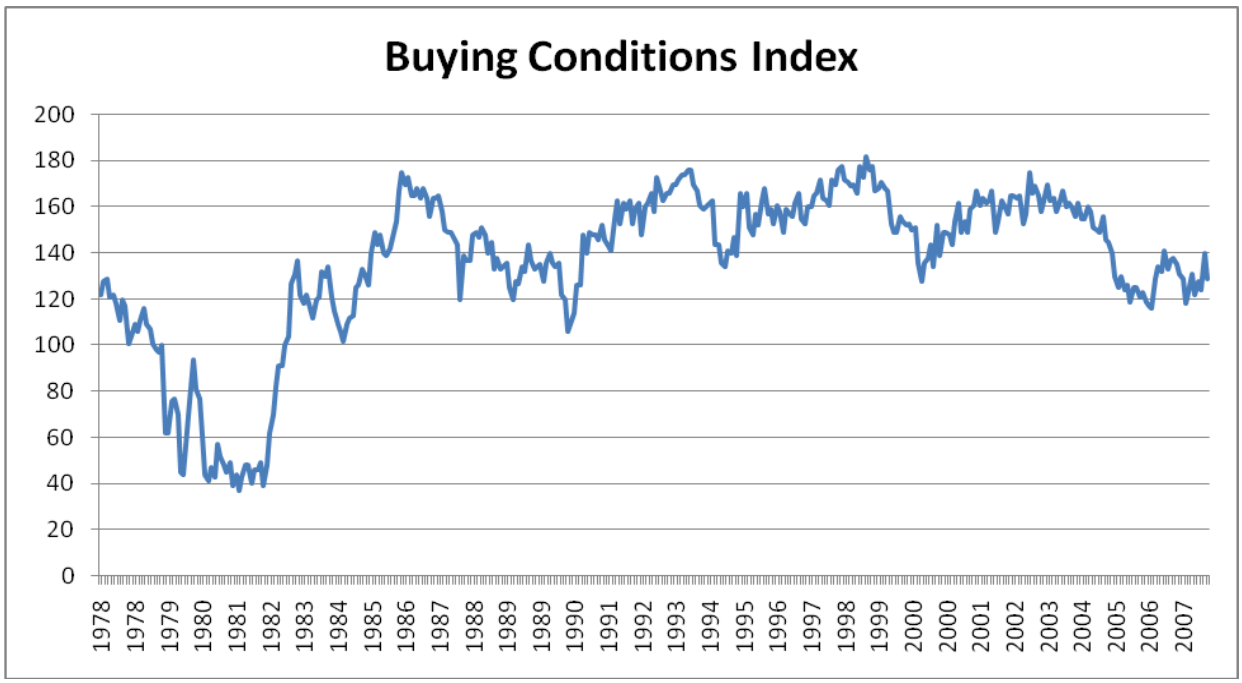
We were able to find this set of questions regarding buying conditions dating all the way back to January 1978, but selling conditions data was only available starting in November 1992. The most recent data available for both of these data sets was March 2008. Therefore, when we look at buying data alone, we use data back to 1978, but when we compare buying and selling data together, we use data going back only to 1992.

Analysis of Summary Statistics: Buying and Selling Sentiment Indexes

Summary of Home Buying Index (1978-2008)

Variable	Observations	Mean	Standard Deviation	Min.	Max
Home Buying Index	362	136.1961	33.53896	37	182

One might expect that when consumers answer the question of whether it's a good or bad time to buy a house, they are comparing current conditions some sort of market average. In this case, we would regularly expect that half of the time, people would respond that it is a relatively "good" time to buy, and the other half of the time, they would say it is a relatively "bad" time to buy. If this were the case, we would expect the mean of the buying index to equal 100, but this is not the case at all. As you can see from the Summary of the Home Buying Index above, the mean index over the thirty-year period is about 136. To test whether our average index value was statistically significantly different from 100, we constructed a 95% confidence interval and saw that it was (132.6, 139.6) (see Appendix Table 1.A for t-test). This finding basically states that, since 1978, the number of people who think it's a good time to buy a house outnumber those who think it's a bad time to buy a house by 36 percentage points on average, as shown below.



For purposes of comparison, we also wanted to include the buying index since 1992. We include it below:

Summary of Home Buying Index 1992-2008 (All Respondents)

Variable	Observations	Mean	Standard Deviation	Min.	Max
Home Buying Index	185	153.9568	15.50169	116	182

Our finding for buying sentiment from 1978-2008 is confirmed for buying sentiment from 1992-2008. The buying index is significantly greater than 100 (see Appendix Table 6.A).

We also took a look at the home selling index to see how it compared to buying sentiment. The data for selling did not go back as far as 1978, so we looked at the years 1992-2008.

Summary of Home Selling Index 1992-2008 (All Respondents)

Variable	Observations	Mean	Standard Deviation	Min.	Max
Home Selling Index	185	118.9189	32.22136	14	164

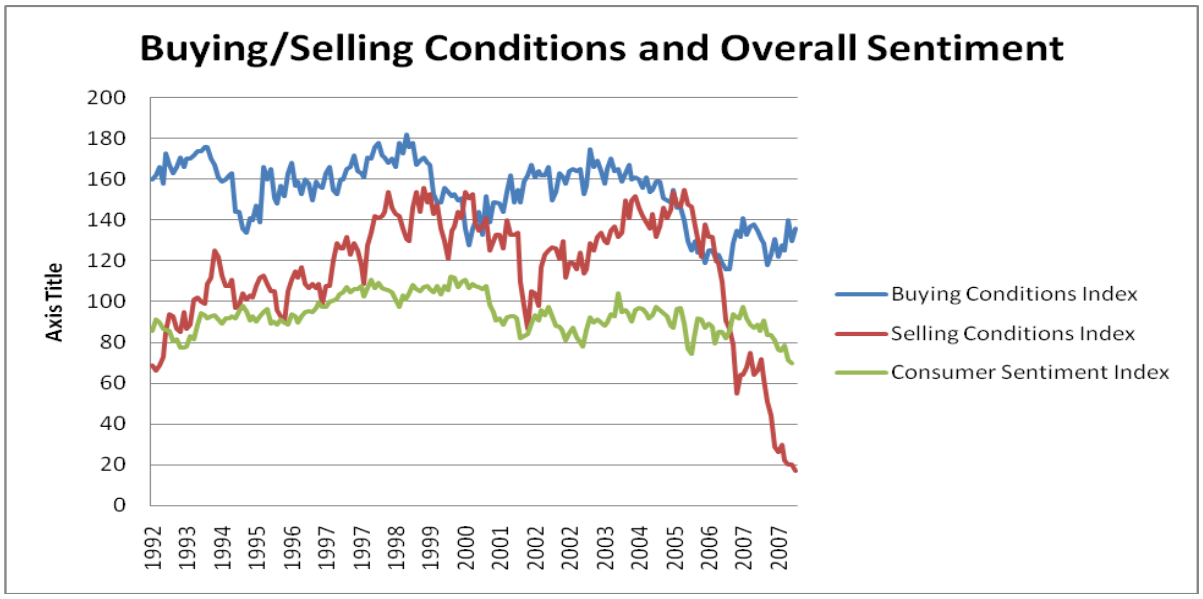
As in the case of our buying statistic, we found a considerable consumer optimism regarding selling conditions over the 15-year span we studied. The mean of the selling conditions index appeared to be considerably higher than 100, so we constructed a 95% confidence interval to assure the statistical

significance of this finding and found that it was indeed significant (114.7, 123.8) (see Appendix Table 11.A for t-test).

Co-Movement of Buying and Selling Sentiment

Analysis of Co-Movement

The fact that buying and selling sentiment are both overwhelmingly positive overtime raises an obvious question: why is it that consumers seem to think that it is simultaneously a good time to buy and sell a house? One would think that a certain set of conditions would favor the buyer, while another would favor the seller, so that there would be some perception of a zero-sum game, but this is clearly not the case. The graph below shows the relative levels of the buying and selling conditions index in the housing market.



Correlation Statistics (1992 to present)

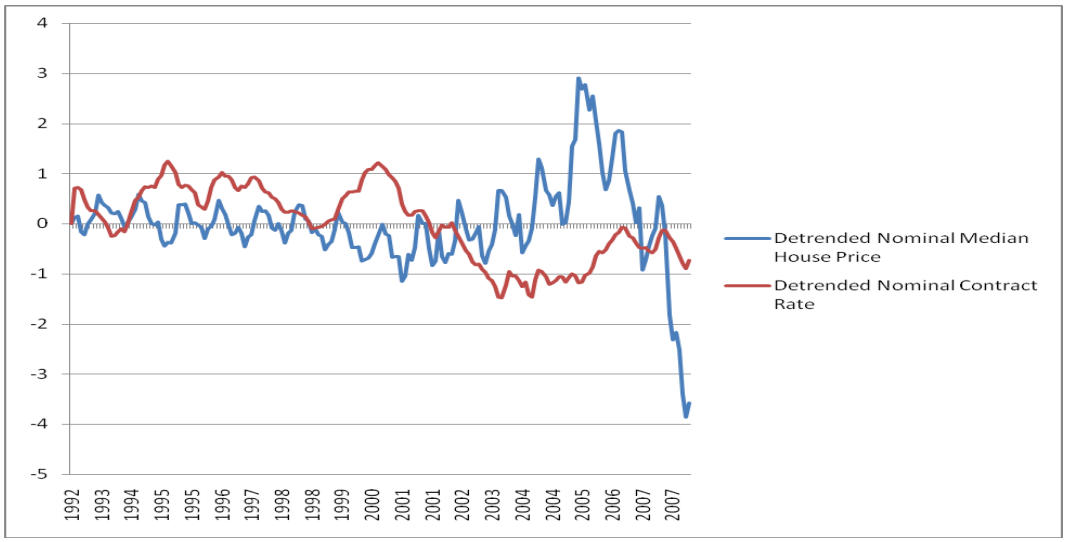
Buying Conditions and Selling Conditions	Buying Conditions and Overall Sentiment	Selling Conditions and Overall Sentiment
0.32300453	0.316826576	0.569059455

Interestingly, selling conditions seem to drive (or be driven by) overall consumer sentiment significantly more than buying conditions. We re-visit this phenomenon later in the paper, but for the moment, we focus primarily on the co-movement of the buying and selling conditions variable.

In line with it being a zero-sum game, our initial prediction was that buying and selling conditions would have a negative correlation, but not necessarily -1. It seems natural to assume a coefficient of slope of -1 when regressing buying and selling conditions because if one additional percent of people thinks it's a good time to buy a house, they would also think it a bad time to sell a house. However, this clearly does not play out in the data, as shown in the table above.

Reasons for Positive Correlation (Prices, Interest Rates, and Inventory/Sales)

One contributing factor to this phenomenon is that some of the reasons given for good buying and selling conditions overlap. For example, "low interest rates" is listed as a reason that it is a good time both to buy and sell a house, as is "good times ahead." Similarly, "high interest rates" and "bad times ahead" are both listed as reasons it is a bad time to both buy and sell a house. If respondents make their assessment of conditions based on a predetermined reason (this assumption is discussed and challenged later), then those who pick one of these overlapping reasons (i.e. "low interest rates" or "bad times ahead") will respond that the market is good (or bad) for both buying and selling. The identical use of interest rates for each sentiment could be explained as follows: lower interest rates make home loans more affordable. This shifts out the demand for housing. An increase in demand necessarily increases the price of houses, which makes it a better time to sell. This hypothesis is confirmed in the graph below, which shows that a decrease in interest rates leads an increase in price and that an increase in interest rates leads a decrease in prices.



The positive correlation may also be explained by business cycles. Over longer periods of time, during booms and recessions, we would expect the two indexes to trend together because they are both indicators of economic well-being (as evidenced by the "good/bad times ahead" response). In the shorter-term, there might still be a negative correlation between the two because a shock in house prices will have opposite ramifications for buying and selling conditions. This disparity between short- and long-term co-movement prompted us to look at buying and selling conditions over shorter periods of time.

When comparing buying and selling data (via scatterplot), we noticed a few structural breaks in the relationship over time. Therefore, we divided our data from 1992-2008 into four periods that had more well-defined relationships between buying and selling sentiment. We display the correlations below:

1992-1997 (55 months)	1997-2002 (60 months)	2002-2006 (41 months)	2006-2008 (28 months)
-0.0975	0.0173	-0.4441	-0.3852

Therefore, although there is a clearly positive long-term trend, the short-term data clearly shows that

the relationship tends to be either negative or non-existent, depending on the era.

Another potential explanation for the co-movement of buying and selling sentiment is home inventory to sales ratio[1]. We predicted that when this ratio is high, houses are going unsold, leading to the perception that it is both a bad time to buy and sell. We collected inventory to sales data from the U.S. Census Bureau website to run the two regressions listed below: one of the home buying index on the inventory to sales ratio and one of the home selling index on inventory to sales ratio. Looking at the regressions, the sign on the inventory-sales ratio is negative for both indexes; an increase in the inventory-sales ratio causes a decrease in the corresponding home sentiment index. This was consistent with our prediction. One interesting finding from these regressions is that the portion of selling sentiment explained by inventory to sales is much higher than that for buying sentiment—the r-square value is almost double for selling sentiment. The likely explanation is that the inventory-sales ratio is a direct reflection of ability to sell. A high inventory of homes necessarily means it is a bad time to sell because unsold homes are increasing. This likewise implies that it is a bad time to buy, but less directly: a high inventory for buyers could also represents more homes choices and lower prices, two encouraging factors for buying sentiment.

Regression of the Home Buying Index on the Inventory to Sales Ratio:

Home Buying Index = $\beta 0 + \beta_1$ Inventory-Sales Ratio + u

Observations = 185

R^2 = .3605

Independent Variables	Coefficient	Standard Error	T	P-Value
Inventory-Sales Ratio	-6.308444	.6211043	-10.16	0.000
Constant	186.0275	3.287156	56.59	0.000

Regresion of the Home Selling Index on the Inventory to Sales Ratio:

Home Selling Index = $\beta 0 + \beta_1$ Inventory-Sales Ratio + u

Observations = 185

R^2 = .6695

Independent Variables	Coefficient	Standard Error	T	P-Value
Inventory-Sales Ratio	-16.30928	.8469956	-19.26	0.000
Constant	198.4534	4.482671	44.27	0.000

Components of Co-Movement

To better understand the dynamics and underlying factors of the co-movement phenomenon, we decided to analyze the co-movement of the individual determinants (i.e. reasons responses) of overall

[1] Inventory to sales data obtained from the Census Bureau, http://www.census.gov/const/fsalmon.pdf

sentiment to see if we could isolate the main drivers of this pattern. To complete this analysis, however, our group had to make a basic assumption:

(1) People enter the survey with a reason for why they think the market is good or bad, and this reason determines their overall sentiment.

We adopt this position over the logical alternative:

(1*) People take the survey with an underlying assumption of the overall goodness/badness of the market and then attempt to conjure up a reason as to why they feel this way.

This assumption allows us to look at the components of the co-movement as sorts of independent factors driving overall co-movement. We are hoping that these co-movements have a logical structure to them (they follow basic economic theory), so that we may isolate one or two intuitive reasons for the overall co-movement present in the data.

It's worth noting that the order of the questions in the survey, in some sense, forces survey-takers into the latter mode of thinking (statement 1*), which may pollute some of the data. We will assume for the time being that had survey-takers been asked the reasons for their sentiments first, that the overall sentiment statistics would remain largely unchanged (this assumes that survey-takers are somewhat decided in how they feel). This assumption will naturally be weaker in the years where the uncertainty statistic (those who answered neither "good" nor "bad") is relatively high.

Test of Independence of Factors

To test the previous assumption, we constructed a correlation matrix to analyze which reason variables tend to move together.

Correlation Matrix of Buying Response Reasons (1992 to 2008):

	Good- Low Prices	Good- Rising Prices	Good - Low Rates	Good- Rising Rates	Good- Good Investment	Good- Good Times	Bad - High Prices	Bad - High Rates	Bad - Can't Afford	Bad- Bad Times	Bad -Will Lose Money
Good- Low Prices	1.00										
Good- Rising Prices	-.43	1.00									
Good- Low Rates	-.07	-.12	1.00								
Good- Rising Rates	-.27	0.60	-.24	1.00							
Good- Good Investment	-.51	0.40	0.06	0.14	1.00						
Good- Good Times	-.30	0.43	0.15	0.19	-.004	1.00					
Bad- High Prices	-.06	-.005	-.7	0.05	0.12	-.41	1.00				

Bad- High Rates	0.18	-.10	-.85	0.18	-.22	-.32	0.71	1.00			
Bad- Can't Afford	0.43	-.32	-.71	-.06	-.20	-.57	0.67	0.80	1.00		
Bad- Bad Times	0.21	-.36	-.28	-.13	-.22	-.62	0.41	0.38	0.53	1.00	
Bad- Will Lose Money	0.36	-.18	-.61	.013	-.20	-.37	0.54	0.67	0.74	0.37	1.00

We looked at this correlation matrix first intending to determine whether the reasons were independent of one another or whether "good" reasons tended to move together even though they were seemingly unrelated. We figured that if our earlier assumption was true, then seemingly independent factors (such as "high prices" and "rising prices") would demonstrate no correlation, while things we know to be negatively correlated (such as "rising prices" and "rising interest rates") would demonstrate negative correlations even though they both indicate a good time to buy.

It turns out that there is always either an insignificant relationship or a significant movement together of "bad" reasons, and never any kind of significant opposite movement (see the bottom right portion of the matrix). We found something similar for the "good" reasons, except in the case of "low prices" which seems to be correlated with an increase in "bad" sentiment. When looking at the time series, this makes perfect sense because the lowest prices occurred at the most recent times, when buying sentiment was lowest.

In addition, we noticed that the strength of relationships between the "good time to buy" variables was far lower than the "bad time to buy" variables. The highest correlation found in the upper left corner of the triangle (good reasons) is 0.60, which is roughly the average correlation found in the bottom right corner (bad reasons).

Unexpectedly weak correlations: There are some variables we expected to have extremely strong correlations, such as "increasing prices" with "good investment" (0.40), and "high prices" with "low prices" (-0.06), that weren't so strong.

Unexpectedly strong correlations: Oddly enough, some of the strongest positive correlations in the matrix came from what we had earlier conjectured to be negatively correlated. Remember earlier that we showed that interest rates and prices tended to be inversely related because of the shift in demand caused by a movement in the interest rate. The "high prices" and "high interest rates" variables (0.71) as well as "high prices and "low interest rates" (-0.70), and "rising prices" and "rising interest rates" (0.60) were very surprising in both strength and sign.

Interesting relationship: "Can't afford" was considerably more strongly correlated with "high rates" (0.80) than it was with "high prices" (0.67). This is finding is not so surprising given what we found earlier in the paper that suggested that consumers consider interest rate to be a better determinant of housing cost than prices.

The above trends also hold true when looking at the data back to 1978, but the sign on the "low prices" correlations are fairly unique to this time period. However, the results of this correlation matrix throw a very difficult twist in our assumption about consumers taking the survey with a pre-conceived reason as to why conditions were good or bad, and then base their sentiment on this underlying reason. Due to

the strong positive correlations among "bad" reasons with one another, (and "good" with the exception of "low prices"), even in the presence of supposedly opposite-moving factors (prices and interest rates), we are forced to conclude that consumers likely have a stronger sense of their general housing sentiment, but a weak sense of their reason for believing so.

This phenomenon is strong enough to cause opposite-moving factors (prices and interest rates) to appear to move together in the minds of consumers. In other words, if there is a general move toward good sentiment in buying conditions, there will be an increase in both the number of people who think interest rates and prices are low, even though these two measures will rarely ever move together..

Later we will explore how to measure this "consumer sentiment regarding housing price and interest rate" by creating variables to gauge this. But for the time being, it is worth remembering that consumers are more likely to move from a general sentiment to a specific reason rather than vice versa.

Uncertainty

When describing the summary statistics, we noted that our analysis was based on a small majority of the population that responded conclusively about their sentiment and gave at least one of the reasons listed in the table at the beginning. This small majority was somewhat disconcerting, so we wanted to gain some insight as to what caused consumer uncertainty regarding conditions in the housing market. Here uncertainty is defined as the percentage of people who answered neither "good" nor "bad" when asked about buying/selling conditions. We first decided to create an "agreement" statistic defined as the absolute difference between the percentage respondents responding "good" and "bad" when asked about overall buying/selling conditions in the housing market.

Agreement = Abs[(% Good) – (%Bad)]

The reason we call this "agreement" is because, if everyone agrees that the market is either "good" or "bad", this measure will be 100, whereas if there is a 50-50 split (total disagreement), this measure is 0. Therefore, we surmised that this statistic is a good measure of market agreement. We were curious as to how much of consumer uncertainty was predicted by what we called a "neutral market", or a situation where there was very low disparity between good and bad sentiments (i.e. high disagreement).

We predicted that as Agreement increased, then Uncertainty would decrease because it would be very clear to most consumers whether it was a good or bad time to buy a house. Therefore we decided regress the uncertainty statistic on the agreement statistic. Our prediction, therefore, is that we would find a significantly negative correlation between Agreement and Uncertainty.

Regression of Buying Uncertainty on Buying Agreement from 1978 to 2008:

$\text{Uncertainty} = \beta_0 + \beta_1 \text{Agreement} + u$

Observations = 361

$R^2 = .0489$

Independent Variables	Coefficient	Standard Error	t	P-Value
Agreement	-.0239088	.0055675	-4.29	0.000

Constant	5.604817	.2741501	20.44	0.000

Regression of Buying Uncertainty on Buying Agreement from 1992 to 2008:

$\text{Uncertainty} = \beta_0 + \beta_1 \text{Agreement} + u$

Observations = 185

$R^2 = .0315$

Independent Variables	Coefficient	Standard Error	t	P-Value
Agreement	.021229	.0086949	2.44	0.016
Constant	2.811309	.4880269	5.76	0.000

Regression of Selling Uncertainty on Selling Agreement from 1992 to 2008:

$\text{Uncertainty} = \beta_0 + \beta_1 \text{Agreement} + u$

Observations = 185

$R^2 = .0249$

Independent Variables	Coefficient	Standard Error	t	P-Value
Agreement	-.023158	.0107031	-2.16	0.032
Constant	7.516167	.3552452	21.16	0.000

We find this result to be evident in our buying data dating back to 1978, but not in the buying data dating back to 1992. However, the regression on selling data back to 1992 does yield a negative coefficient on disparity. We are not sure why the 1992 buying data does not confirm our hypothesis on uncertainty. Overall, we found that disagreement in housing market sentiments does a very poor job at explaining the amount of uncertainty we find in the market.

We also regressed buying uncertainty on selling uncertainty to see how much of each statistic was due to some "general uncertainty".

Regression of Selling Uncertainty on Buying Uncertainty from 1992 to 2008:

$\text{Selling Uncertainty} = \beta_0 + \beta_1 \text{BuyingUncertainty} + u$

Observations = 185

$R^2 = .4191$

Independent Variables	Coefficient	Standard Error	t	P-Value
Buying Uncertainty	.9054306	.0788013	11.49	0.000
Constant	3.282296	.3441213	9.54	0.000

From this regression we found that about two-fifths of uncertainty in either buying or selling sentiment can be explained by a general sense of uncertainty in the housing market. The other three-fifths is buying- or selling-specific.

II. Consumer Misunderstanding—Determinants of Price and Interest Rate Sentiment

Analysis of Summary Statistics: Buying and Selling Sentiment Reasons

It's worth noting that the data we researched first grouped respondents into the general "good" or "bad" categories, and then evaluated their reasons conditional on which group they were in. It was much more convenient for our research purposes to instead have variables that told us the percentage who replied good or bad combined with the appropriate reason, out of the entire population rather than a particular subgroup. Therefore, we rescaled the data into "Population variables". For example, we created the variable:

```
"buy_pop_good_low prices" = (% good) * (% low prices | good) / 100.
```

The example above shows how, for any given month, we were given the percentage of respondents saying that buying conditions were favorable, as well the percentage of those people who cited low prices as the reason. We wanted to transform this variable so that we could interpret the total percentage of the population who thought it was a good time to buy a house because of low prices, and the formula is shown above.

We performed this rescaling for every variable, and then looked at the summary statistics. The results for the home buying response reasons from 1978 to 2008 are shown below.

Summary of Buying Response Reasons 1978-2008 (All Respondents)

Variable	Observations	Mean	Standard Deviation	Min.	Max
(1) Good- Low Prices	362	13.05616	8.338298	.54	38.64
(2) Good- Rising Prices	362	5.482486	3.619452	.6	22.2
(3) Good- Low Rates	362	29.39196	19.24012.18	.18	68.53
(4) Good- Rising Rates	362	5.035829	3.816781	0	18.2
(5) Good- Good Investment	362	4.348315	1.755616	.45	9.24
(6) Good- Good Times	362	3.251236	2.897664	0	12.18
(7) Bad- High Prices	362	5.172072	5.629905	.28	24.64
(8) Bad- High Rates	362	8.333122	14.54935	.07	65.57
(9) Bad- Can't Afford	362	2.36326	2.292392	.07	13.68

(10) Bad- Bad Times	362	1.075166	1.126029	0	6.84

Population percentages of buying response reasons:
(1) good time to buy: low prices (2) good time to buy: increasing prices (3) good time to buy: low interest rates (4) good time to buy: rising interest rates (5) good time to buy: good investment (6) good time to buy: good times financially (7) bad time to buy: high prices (8) bad time to buy: high interest rate (9) bad time to buy: can't afford (10) bad time to buy: bad times ahead.

Note that the sum of the means here only adds up to 77.5. This basically means that at any point in time, an average of 77.5% of the population had a definitive ("good" or "bad") opinion about the housing market that was based on the 10 reasons that the survey created categorical variables from. It is this 77.5% of the population that we do most of our analysis on.

From the Summary of Buying Response Reasons table above, since 1978, consumer sentiment regarding the favorability of buying conditions in the housing market is based foremost on interest rates, and secondly on prices. We determined this by looking at the relative sizes of the reason variables. To confirm this finding statistically, we ran four separate t-tests (see Appendix Tables 2.A – 5.A). First, we tested that the percentage of people who answered because of interest rates was higher than those who answered because of prices. Then we tested that the prices response was higher than the next highest response. We did this both for the group who responded "good" and the group that responded "bad". Every result was significant at the 1% level.

For comparison, we look at the buying statistics from 1992-2008:

Summary of Buying Response Reasons 1992-2008 (All Respondents)

Variable	Observations	Mean	Standard Deviation	Min.	Max
(1) Good- Low Prices	185	14.19005	7.140878	4.88	38.64
(2) Good- Rising Prices	185	4.816486	1.90458	.64	9.62
(3) Good- Low Rates	185	38.75124	15.43948	4.88	68.53
(4) Good- Rising Rates	185	5.650595	4.206029	0	18.2
(5) Good- Good Investment	185	4.997351	1.622359	1.24	9.12
(6) Good- Good Times	185	4.963027	2.996399	0	12.8
(7) Bad- High Prices	185	2.593784	2.440873	.28	11.7
(8) Bad- High Rates	185	1.927189	1.975157	.07	8.14
(9) Bad- Can't Afford	185	1.534216	1.365094	.07	6.93
(10) Bad- Bad Times	185	.5278378	.3665463	0	2.07
(11) Bad- Lose Money	185	.0953514	.1539657	0	.74

Population percentages of buying response reasons:
(1) good time to buy: low prices (2) good time to buy: increasing prices (3) good time to buy: low interest rates (4) good time to buy: rising interest rates (5) good time to buy: good investment (6) good time to buy: good times financially (7) bad time to

buy: high prices (8) bad time to buy: high interest rate (9) bad time to buy: can't afford (10) bad time to buy: bad times ahead.

Our finding is confirmed for this 16 year period: consumer sentiment for buying conditions is based primarily on interest rates and prices (see Appendix Tables 7.A – 10.A for t-test confirmation).

We then look at the response reasons for home selling conditions:

Summary of Selling Response Reasons 1992-2008 (All Respondents)

Variable	Observations	Mean	Standard Deviation	Min.	Max
(1) Good- High Prices	185	9.518541	5.488996	.12	25.74
(2) Good- Falling Prices	185	1.01373	.8243005	0	4.38
(3) Good- Low Rates	185	12.03103	6.195115	.08	24.96
(4) Good- Rising Rates	185	1.150811	1.134632	0	4.5
(5) Good- Good Investment	185	2.289892	1.520742	0	8.14
(6) Good- Good Times	185	9.114108	5.095135	0	19.5
(7) Bad- Low Prices	185	9.544378	10.90258	1.14	55.2
(8) Bad- High Rates	185	2.367189	2.533822	.17	14.62
(9) Bad- Can't Afford	185	4.676973	5.051313	.28	28.16
(10) Bad- Bad Times	185	1.038703	1.109772	0	6.3
(11) Bad- Lose Money	185	1.902108	2.286148	0	13

Population percentages of selling response reasons:
(1) good time to sell: high prices (2) good time to sell: decreasing prices (3) good time to sell: low interest rates (4) good time to sell: rising interest rates (5) good time to sell: good investment (6) good time to sell: good times financially (7) bad time to sell: low prices (8) bad time to sell: high interest rate (9) bad time to sell: can't afford (10) bad time to sell: bad times ahead (11) bad time to sell: lose money.

Here we found that consumers also base their sentiment of selling conditions primarily on interest rates and prices (see Appendix Tables 12.A – 15.A).

Note that the sum of the means for selling here only adds up to 54.6. This basically means that at any point in time, less than 54.6% of the population had a definitive ("good" or "bad") opinion about the housing selling market that was based on the 11 reasons that the survey created categorical variables from. We say "less than" because respondents were allowed to list multiple reasons, allowing for double-counting among these reasons. Therefore, it is on this 54.6% (or less) of the population that we do most of our selling conditions analysis.

Implications of Summary Buying and Selling Statistics

Once we statistically confirmed that interest rates and then prices are most important in determining consumer sentiment, we sought out a reason as to why consumers would place a sort of importance premium on interest rates over prices for buying conditions sentiment. Certainly both have a strong influence on the ultimate cost or affordability of housing. It seemed to us that consumers would tend to shy away from basing their opinion on the state of the market based a relatively stable measure of cost. It seemed more intuitive that consumers would base their opinions on something in the market that fluctuated considerably with differing economic conditions. Therefore, we hypothesized that the tendency to base sentiments on a certain measure reflected a general sense of perceived volatility of that measure by the consumer. As a result, we concluded that consumers likely consider interest rates to be a more volatile measure of housing cost than the house price itself.

To analyze the validity of this perception, we looked at the standard deviations of the national average contract mortgage rate and median real house price relative to their means. We only looked at data from 1992 to the present, since this is when we had data on both buying and selling sentiments, and these measures together were the basis of our last conclusion. Here is what we found:

Summary of Contract Rate and Real Median House Price 1992-2008

Variable	Observations	Mean	Standard Deviation	Min.	Max	(Std. Dev./Mean)
Contract Rate	185	6.835351	.7271557	5.36	8.08	.1063816
Real Median House Price	185	126370.1	17853.49	101827.3	167118.6	.1412794

When we calculated the standard deviation relative to the mean for each rate, we found that the contract rate had a scaled standard deviation of 0.106 “means” and median house price had a scaled standard deviation of 0.141 “means”. In this sense, we found that, contrary to the suggestion of the relative importance of price and interest rates, house prices tend to be more volatile than interest rates according to this measure. This could be the first instance of consumer misunderstanding regarding the housing market.

Perceived Prices and Interest Rates

It became apparent very quickly that the two main components that influenced buying and selling sentiment in regards to the housing market were house prices and interest rates. It was clear that the perception of these factors was the main driving force in housing market sentiment. In order to compare perceptions against actual levels, however, we needed a variable that represented consumer perception of these two quantities, or at least determined how consumer attitudes indicated their perception of them.

It must be stressed that, although we used names like “Perceived Price” and “Perceived Interest Rate” for our variables, these are both significant misnomers. These variables do not measure how all consumers perceive the price level or interest rate level, but only how consumers are basing their general housing sentiments on prices and interest rates. For example, a consumer may think that prices

are high, but that interest rates are so low that she still considers it a good time to buy a house. This person will only be counted in the "low interest rate" category, and not the "high price" category, since "high price" did not determine her overall sentiment. Therefore, there is a considerable lack of information if one chooses to interpret these as the actual perceived price or interest rate level, rather than how prices and interest rates are actually affecting housing sentiment.

We created the variables according to the following formulas:

Perceived interest rate = [(Hb + Hs) - (Lb + Ls)]/2 + 100

(1)Lb = % responding that is good time to buy a house because of low interest rates
(2)Hb = % responding that is a bad time to buy a house because of high interest rates
(3)Ls = % responding that is good time to sell a house because of low interest rates
(4)Hs = % responding that is bad time to sell a house because of high interest rates

Perceived Price = [(Hb + Hs) - (Lb + Ls)]/2 + 100

(1)Lb = % responding that is good time to buy a house because of low prices
(2)Hb = % responding that is a bad time to buy a house because of high prices
(3)Ls = % responding that is bad time to sell a house because of low prices
(4)Hs = % responding that is good time to sell a house because of high prices

Naturally, we were only able to construct these variables as such since 1992 because selling information was only available after this time. Like our other index variables, these have a potential range of 0-200 and take into account the level of prices and interest rates people perceive both in buying and selling. Below is a summary of these index variables and their components:

Summary of Perceived Price

Variable	Observations	Mean	Standard Deviation	Min.	Max
Perceived Price	185	94.12459	11.464	54.245	113.09

Summary of Perceived Price Components

Variable	Observations	Mean	Standard Deviation	Min.	Max
Selling- Good-High Prices	185	10.17703	5.729338	.08	26.52
Selling- Bad-Low Prices	185	9.531676	11.18047	.76	54.9
Buying- Good-Low Prices	185	14.72503	7.450776	4.96	39.2
Buying- Bad-High Prices	185	2.328865	2.250377	.24	9.9

(1) Population variable of those who say good time to sell because of high prices (2) Population variable of those who say bad time to sell because of low prices (3) Population variable of those who say good time to buy because of low prices (4) Population variable of those who say bad time to buy because of high prices.

Summary of Perceived Interest Rate

Variable	Observations	Mean	Standard Deviation	Min.	Max
Perceived Rate	185	75.18843	12.55559	54.065	104.44

Summary of Perceived Interest Rate Components

Variable	Observations	Mean	Standard Deviation	Min.	Max
Buying- Good- Low Interest Rates	185	40.14978	15.93903	4.96	69.3
Buying- Bad- High Interest Rates	185	1.754757	1.838939	.06	7.48
Selling- Good- Low Interest Rates	185	13.75049	7.088292	.08	29.64
Selling- Bad- High Interest Rates	185	2.382378	2.560822	.17	14.62

(1) Population variable of those who say good time to buy because of low interest rates (2) Population variable of those who say it is a bad time to buy because of high interest rates (3) Population variable of those who say good time to sell because of low interest rates (4) Population variable of those who say it is a bad time to sell because of high interest rates

Though we wanted to include all reason variables to construct our perceived interest rate for a more holistic view, note that the variables representing the percentage of individuals who think it's a good time to buy because of low interest rates and those who think it's a good time to sell because of low rates dominate this index measure (meaning it is almost always less than 100). Therefore, our perceived interest rate will be mostly a reflection of these two variables.

Consumer Compensation: Compounding Interest

We were afforded the option of using either the "contract" rate of interest (that which is visible on the loan itself) or the effective rate of interest in measuring the actual interest rate on housing[2]. The effective rate of interest seems to be an indicator of the final cost of a home because it accounts for compounding effects, but we were curious as to whether consumers based their decisions more on the contract rate because it is more visible. We decided to regress perceived interest rate on both of these measures to see if effective rate became a pretty useless predictor in the presence of contract rate (i.e. consumers pay no attention to compounding effects). We noticed a significant structural break in interest rate trends (more on this later), so we broke up our regression into two smaller pieces where the trends were fairly consistent. The first period is 11/92 to 8/02, and the second is 9/02 to 3/08.

Regression of Perceived Interest Rate on Effective and Contract Rate from November 1992 to August 2002:

Perceived Interest Rate $= \beta_0 + \beta_1$ Effective Rate $+ \beta_2$ Contract Rate $+ u$

Observations = 118

$R^2 = .7403$

Independent Variables	Coefficient	Standard Error	T	P-Value

[2] Contract interest rate data obtained from the Federal Reserve Board and effective interest rate data obtained from the National Association of Realtors.

Effective Rate	-30.88293	12.76118	-2.42	0.017
Contract Rate	50.95585	13.47189	3.78	0.000
Constant	-67.68322	8.067558	-8.39	0.000

Regression of Perceived Interest Rate on Effective and Contract Rate from September 2002 to March 2008:

Perceived Interest Rate = $\beta_0 + \beta_1$ Effective Rate + β_2 Contract Rate + u

Observations = 67

R^2 = .7091

Independent Variables	Coefficient	Standard Error	T	P-Value
Effective Rate	-32.46262	37.8363	-0.86	0.394
Contract Rate	68.35785	39.7895	1.76	0.083
Constant	-135.5536	17.34119	-7.82	0.000

We expected a positive coefficient on the contract rate (the rate that homebuyers actually see), and an insignificant coefficient on the effective interest rate (the rate the homebuyers actually pay). This hypothesis was correct for the most recent period, but there was actually a significant negative coefficient on the effective rate before 2002. This leads us to conclude that homebuyers do not compensate for compounding interest in their perceptions of interest rates on housing, and often compensate "backwards" for it.

This finding is important in and of itself, but it was also important in selecting which variable to use in order to maximize explanatory power in later regressions. We decided to use contract rate in each case because it more clearly represented the interest rates consumers are considering when they formulate their housing market sentiments. We were actually interested in seeing how much extra explanatory power the contract rate afforded us, so we decided to regress perceived interest rate on each variable and check how the R-squared statistics compare. In both time periods, contract rate provides a better fit and adds 1-2% explanatory power.

This has quite a few implications for banks, and highly encourages banks to use a simple interest rate so that they can make their contract rate as low as possible while providing the same effective rate as other banks.

Consumer Compensation: Inflation

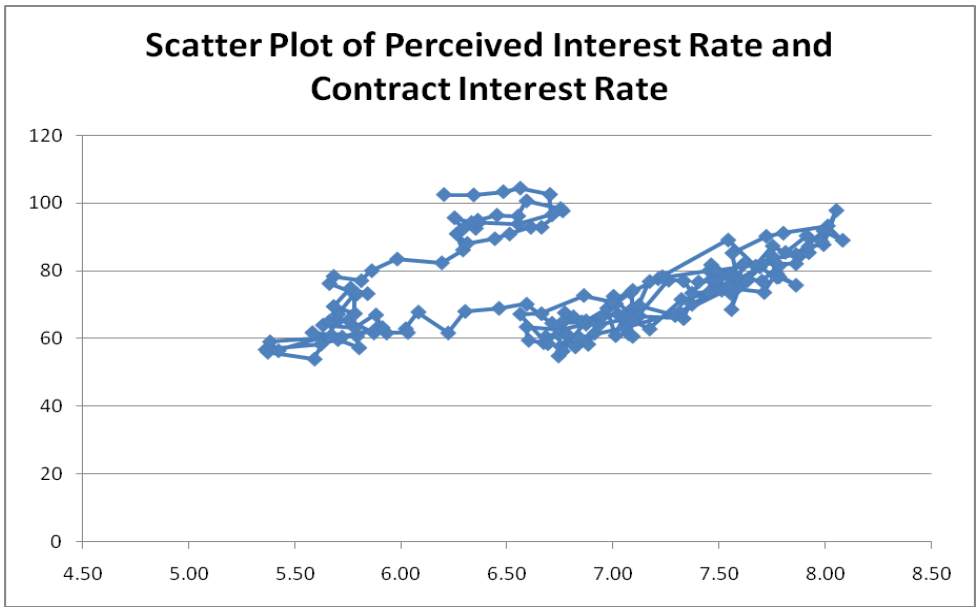
In view of our large-scale goal of determining consumer sentiment in the housing market, we discovered that house price and interest rate were the two dominating factors. It was also clear that both of these measures were quite influenced by the rate of inflation, and we wanted to see how good consumers were at discerning (and accounting for) the rate of inflation[3] as it affects these two measures. We therefore embarked on comparing two regressions:

[3] Inflation data obtained from Bureau of Labor Statistics

(1) Perceived Interest Rate = $\beta_0 + \beta_1$Contract Rate + u

(2) Perceived Interest Rate = $\beta_0 + \beta_1$ Real Rate + u

"Real rate" is simply the contract rate adjusted for inflation. However, when we ran both regressions over the time period 1992-2008, we disappointingly had a very poor fit ($R^2 = 0.07$). We decided to investigate why this was the case, and found that there was a very clear structural break in the interest rate data, as found below:



We saw a very clear linear trend in the "lower" leg on the very right (when interest rates were 6.5-8.5%), followed by a short horizontal segment, into the "upper leg" (in a time period when interest rates were 5-6.5%). As a result, we decided to split the regression into two separate pieces to account for each of these separate trends. It turns out that this lower leg occurred between November 1992 and August 2002 (118 months). The transition period and upper leg occurred between September 2002 and March 2008 (67 months).

We then ended up with far more satisfactory R^2 values:

Regression of Perceived Interest Rate on Contract Rate (11/92-8/02):

Perceived Interest Rate = $\beta_0 + \beta_1$Contract Rate + u
Observations = 118
R^2 = .7271

Independent Variables	Coefficient	Standard Error	T	P-Value
Contract Rate	18.448	1.049359	17.58	0.000
Constant	-60.51826	7.660164	-7.90	0.000

Regression of Perceived Interest Rate on Contract Rate (9/02-3/08):

Perceived Interest Rate = $\beta_0 + \beta_1$Contract Rate + u
Observations = 67
$R^2 = .7057$

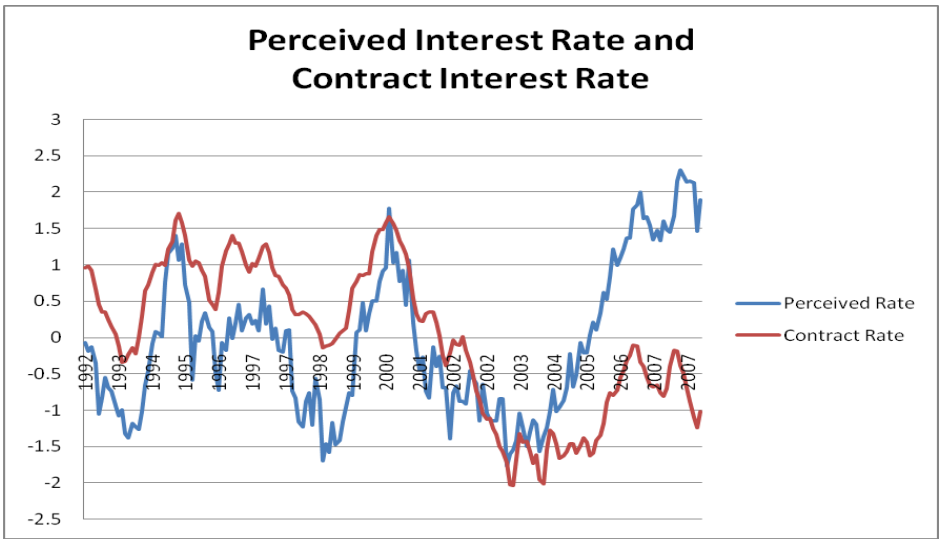
Independent Variables	Coefficient	Standard Error	t	P-Value
Contract Rate	35.16561	2.816543	12.49	0.000
Constant	-132.9647	17.04195	-7.80	0.000

We got a slightly lower R^2 in the second period because it included the horizontal transition period, but because the first regression had such a large number of observations and a very clear trend, we decided to state its R^2 value (slightly under 73%) as a finding: about three-fourths of the movement in the contract interest rate is translated into changing housing market sentiment based on the interest rate.

The second compelling finding from this graph and subsequent regressions was that there was a considerably steeper slope in the "upper leg" of the graph. Not only did an increased number of consumers base their housing sentiments on the presence of a "high interest rate" when the interest was at historically low levels, but they were also more sensitive to changes in the contract rate (as witnessed by a near-doubling of the coefficient from 18.4 to 35.2).

The interpretation here is that a one-percentage-point increase resulted in 18% more of the population basing their housing sentiment on the presence of a high interest rate (as opposed to a low one) in the earlier period, whereas that same change result in 35% more of the population basing their housing sentiment on the presence of a high interest rate in the later period. An easy way to summarize this is that consumers about doubled in sensitivity beginning in 2003 (when we exclude the transition period, our second period essentially begins in 2003).

This finding is illustrated in the chart below, which shows that beginning around 2003, the perceived rate moved at a more drastic slope than the contract rate.



Our search for the effects of inflation on consumer sentiment regarding interest rate certainly produced some interesting findings, but these have yet to address the initial question regarding how consumers take inflation into consideration in determining the interest rate.

Therefore, we decided to perform two new regressions:

(1) Perceived Interest Rate = $\beta_0 + \beta_1$ Contract Rate + β_2 Inflation + u (11/1992-08/2002)
(2) Perceived Interest Rate = $\beta_0 + \beta_1$ Contract Rate + β_2 Inflation + u (09/2002-03/2008)

Our hypothesis was that we would get a negative coefficient on inflation ($\beta_2 < 0$) because, if the inflation rate was higher, consumers would realize that the real interest rate was lower, and so the perceived interest rate would drop (holding contract rate constant, of course). What we found, however, was quite surprising.

In both regressions, $\beta_2 > 0$ at the 10% confidence level (see Appendix Tables 1.C and 2.C). This result was quite surprising, especially considering the cleanness of the trends in each period. It's worth noting that the coefficient on inflation was far more significant (t = 4.21) in the second period than it was in the first (t = 1.92). Since the coefficient in the first period is significant at the 10% level, but not the 5% level, and because we believe it to be the more representative of the regressions, we are hesitant to go all the way to infer that consumers judge inflation backward. For the time being, we infer only that inflation is not correctly taken into account.

In addition, we found that when we regressed perceived rate on the real contract rate of interest by itself, we found an insignificant coefficient on the real interest rate as well as an $R^2 < 0.01$ in both periods (see Appendix Tables 3.C and 4.C). No matter the technique employed (we even tried accounting for other factors we think affects what consumers think about the interest rate, including house price and indicators of general economic well-being including unemployment and the overall Index of Consumer Confidence, as shown below), we could not obtain a negative coefficient on inflation, and most commonly found a statistically-significant positive coefficient. This spoke volumes regarding consumers' ability to properly account for inflation.

Regression of Perceived Interest rate on Contract Rate, Nominal Median House Price, Inflation, the Index of Consumer Sentiment and Unemployment:

Perceived Interest Rate = $\beta_0 + \beta_1$ Contract Rate + β_2 Nominal Median House Price (in thousands of dollars) + β_3 Inflation + β_4 Index of Consumer Sentiment + β_5 Unemployment + u

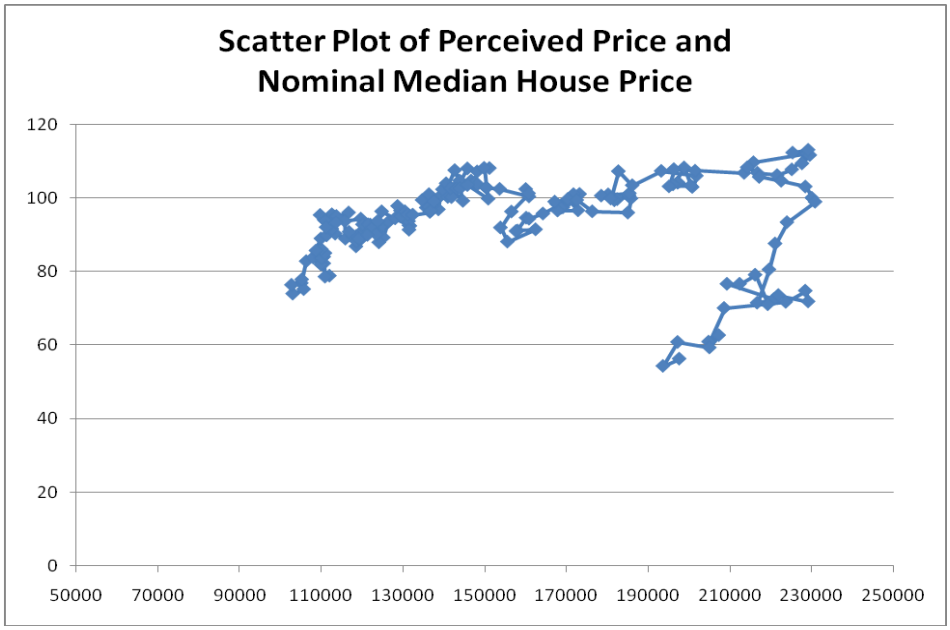
Observations = 185

$R^2 = .7735$

Independent Variables	Coefficient	Standard Error	t	P-Value
Contract Rate	15.51674	1.391789	11.15	0.000
Nominal Median House Price	.2754013	.0300359	9.17	0.000
Inflation	4.188877	.6659337	6.29	0.000
Index of Consumer Sentiment	-.4407659	.0893985	-4.93	0.000
Unemployment	-3.045603	1.353965	-2.25	0.026
Constant	-27.67659	25.57086	-1.08	0.281

The positive coefficient of inflation in these three regressions leads us to conclude that consumers do not properly account for inflation in their evaluation and subsequent sentimient in the housing market regarding interest rates.

We went through a similar process to see how inflation affected consumer sentiment regarding house prices. We first ran a wholesale regression of the perceived price rate on the median nominal house price for the 1992-2008 era and once more discovered a poor R^2. We looked at the scatterplot for structural breaks and this time found two breaks instead of one, as we had for interest rates. We show the scatterplot below.



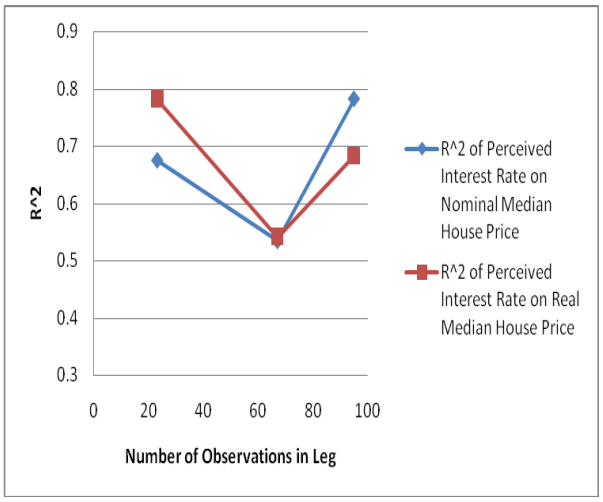
We saw three distinct legs in this scatterplot corresponding to three positive-sloping linear trends (as we had expected) with distinctly different slopes and intercepts. The structural breaks resulted in our running three separate regressions over the following time periods:

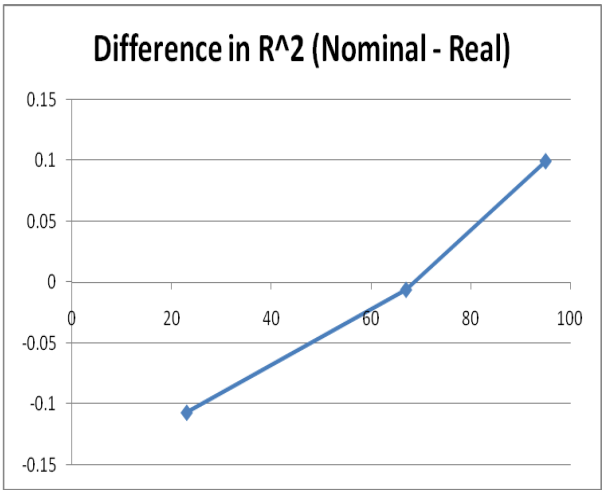
(1) November 1992 – September 2000 (95 months)
(2) October 2000 – April 2006 (67 months)
(3) May 2006 – March 2008 (23 months)

We then performed the following regressions on each period:

(1) Perceived Price = $\beta_0 + \beta_1$ Nominal Median House Price
(2) Perceived Price = $\beta_0 + \beta_1$ Real Median House Price

We found that as the length of the time period increased (i.e. the number of data points), then so did the fit of the regression on the nominal house price relative to that on the real price (see Appendix Tables 1.D – 6.D).





The fit for the first and third legs were considerably better than that for the second. Over the first and third legs, we achieved R^2 values around 75%, whereas in the second period, they were about 55%. We expected nominal prices to be the better predictor in the short-run, but for real prices to be the best long-run predictor, but the graph on the right directly contradicts this idea. This seems to indicate that consumers also poorly compensate for inflation when formulating their sentiment on house prices. To confirm this notion further, we ran the following regressions:

(1) Perceived Price = $\beta_0 + \beta_1$ Nominal Median House Price + β_2 Inflation + u (11/1992-09/2000)
(2) Perceived Price = $\beta_0 + \beta_1$ Nominal Median House Price + β_2 Inflation + u (10/2000-04/2006)
(3) Perceived Price = $\beta_0 + \beta_1$ Nominal Median House Price + β_2 Inflation + u (05/2006-03/2008)

In each case, we got an R^2 value around 0.75 (see Appendix 7.D – 9.D). So about three-fourths of the movement in nominal median house price (and inflation) is translated into changing housing market sentiment based on the price of housing. However, just as was the case for interest rates, there was a positive coefficient on inflation ($\beta_2 > 0$ in every regression). Even with nominal prices held constant, consumers tended to look at inflation backwards. This led us to conclude that consumers do not properly accoutn for inflation in their evaluation and subsequent sentiment in the housing makrety regarding house prices.

We also noticed that β_1 (which we have labeled price sensitivity) fluctuated considerably over the three periods (0.51, 0.11, and 1.04 respectively). This led us to our conclude that consumer sensitivity to prices has undergone sever deviations—first hyposensitivity, then hypersensitivity—since 2001.

III. Moving Toward a Unified Model

All of the structural breaks in both price and interest rate sensitivity led our group to believe that there were other factors at play in a very strong sense that were affecting how consumers thought of the current housing market. These structural breaks were beginning to get a bit annoying, so we decided to move toward a more uniform model that could explain consumer sentiment for the entire 15-year period from 1992 to 2008.

To begin the process of variable selection, we decided it was best to start with the reasons that were most listed in the survey as determinants of consumer sentiment. Our general assumption was that the reasons people gave behind their sentiment actually had backing in real economic conditions. Therefore, we sought real economic variables to represent reasons like "high interest rates", "falling prices", "can't afford", and "bad times ahead". The table below shows the reasons variables that were collected with the survey as well as our real-world proxy:

Survey Response Reason	**Corresponding Economic Indicator**
Low/High Prices	Nominal House Price, Inflation
Rising/Falling Prices	De-trended House Price
Low/High Interest Rates	Contract Rate, Inflation
Rising Interest Rates	ΔContract Rate
Good Investment	De-trended House Price
Good/Bad Times Ahead	Consumer Expectation Sentiment (ICE), ΔUnemployment
Can't Afford	Current Consumer Sentiment (ICC), Unemployment
Will Lose Money	De-trended House Price

We then created saturated models for buying and selling sentiment with each of these variables and performed a backward elimination model selection process to arrive at the most useful models in the end. We used the Bonferroni correction because of our large models and decided to only select variables with a p-value < 0.01.

Considering the very small R^2 values we got by looking at interest rates and prices alone over this entire period, the fits we got in these holistic models were quite satisfying. It turns out that the vast majority of the structural breakage we saw in our perceived price and interest rate regressions can be explained by these real-world proxy variables.

Regression of Home Buying Index on Contract Interest Rate, Nominal Median House Price in thousands of dollars, De-trended House Price in thousands of dollars, Inflation, Unemployment Level, Quarterly Percentage Change in Contract Interest Rate, Yearly Percentage Change in Unemployment Level, and the Index of Consumer Expectations from 11/1992 to 3/2008:

$$\text{Home Buying Index} = \beta_0 + \beta_1 \text{Contract Rate} + \beta_2 \text{NominalHousePrice} + \beta_3 \text{DetrendedHousePrice} + \beta_4 \text{Inflation} + \beta_5 \text{Unemployment} + \beta_6 \text{ContractRateChange} + \beta_7 \text{UnemploymentChange} + \beta_8 \text{ConsumerExpectations} + u$$

Observations = 173
R^2 = 0.8695

Independent Variables	Coefficient	Standard Error	T	P-Value
Contract Rate	-21.75486	1.390742	-15.64	0.000
Nominal House Price	-.5205197	.03083	-16.88	0.000
De-trended House Price	.2834122	.0603047	4.70	0.000
Inflation	-2.658266	.7823252	-3.40	0.001
Unemployment	-3.31794	1.190715	-2.79	0.006
Contract Rate Change	35.71694	13.36007	2.67	0.008
Unemployment Change	20.55295	4.586778	4.48	0.000

Consumer Expectations	.3914198	.0704713	5.55	0.000
Constant	375.5732	21.90125	17.15	0.000

Regression of Home Selling Index on Contract Interest Rate, Nominal Median House Price in thousands of dollars, De-trended House Price in thousands of dollars, Inflation, Unemployment Level, Quarterly Percentage Change in Contract Interest Rate, Yearly Percentage Change in Unemployment Level, and the Index of Consumer Expectations from 11/1992 to 3/2008:

Home Selling Index $= \beta_0 + \beta_1$ Contract Rate $+ \beta_2$ NominalHousePrice $+ \beta_3$ DetrendedHousePrice $+ \beta_4$ Inflation $+ \beta_5$Unemployment $+ \beta_6$ ContractRateChange $+ \beta_7$UnemploymentChange $+ \beta_8$ConsumerExpectations $+ u$

Observations = 173
$R^2 = 0.8067$

Independent Variables	Coefficient	Standard Error	T	P-Value
Contract Rate	-39.74705	3.129876	-12.70	0.000
Nominal House Price	-.7887699	.0693831	-11.37	0.000
De-trended House Price	2.26503	.1357163	16.69	0.000
Inflation	12.06819	1.760629	6.85	0.000
Unemployment	-20.20593	2.679713	-7.54	0.000
Contract Rate Change	131.4703	30.06694	4.37	0.000
Unemployment Change	72.20845	10.32258	7.00	0.000
Consumer Expectations	1.469863	.1585962	9.27	0.000
Constant	461.3129	49.28893	9.36	0.000

Interestingly enough, even when starting with very large saturated models, we arrived at identical final models. It was quite exciting to explain 87% and 81% of buying and selling sentiment respectively over such a diverse and volatile period. Quite interestingly, every predictor except inflation has the same sign in each regression, showing a fairly similar and consistent formulation of buying and selling sentiment, even though we were expecting opposite signs for the house price variables.

We wondered if these similarities in coefficients were simply because buying and selling sentiment moved together over this period, but when we included the buying statistic in the selling regression and vice versa, we were met with insigifnicant results.

The results of these phenomena are a bit difficult to interpret. It seemed to us that the inflation coefficient should be positive in each regression because higher inflation meant a lower real rate of interest, which would result in people thinking it a good time to buy and sell a house. However, we only found this in the selling data. Also, it is still unclear as to why the house price variables show the same sign on the coefficeint in both the buying and selling data, though it is heartening to see that de-trended house price was fare more meaningful in selling sentiment, because it seems that people with opinions on this matter (most likely homeowners) are clearly informed about prices relative to normal levels, and the sign for this was correctly positive.

IV. Implications: The Recent Housing Crash

Homeowners vs. Non-homeowners

Beginning in 1992, the data on buying and selling sentiment is split into two categories: all respondents and homeowners. We thought it would be interesting to distinguish between the perceptions of homeowners and non-homeowners on the market, because this paper is primarily interested in discovering the determinants of consumer sentiment. Our hypothesis coming in was that we would see significantly more disparities in the selling sentiments between homeowners and non-homeowners than we would for buying because everyone in the market is a potential buyer, but only *homeowners* are potential sellers. For this reason, we predicted that homeowner selling sentiment was considerably more informed and would show a closer relationship to actual market conditions.

Note on the data: Since the data was originally split up into homeowners and all respondents, we had to formulate for ourselves the data on the non-homeowners category. We used the following formula:

All = p_{HOME}*Home+ $p_{NONHOME}$*Nonhome
Where p_{HOME} = the proportion of respondents that are homeowners
And $p_{NONHOME}$ = the proportion of respondents that are not homeowners
(both of these variables were easy to calculate because we were given both the number of homeowners and all respondents for each month)

From this, we solved for and constructed the "Non-home" variable using the "All" and "Home" variables. It's worth noting that the "All" and "Home" variables are all integer-valued, but because of this formula we used, the "Non-home" data is not. We decided to leave it this way rather than round because we felt that it provided us the most accurate information, even if it was inconsistent with the format of the others.

We first compared the indexes of homeowners and non-homeowners for buying data:

Summary of Home Buying Index 1992-2008 (Homeowners)

Variable	Observations	Mean	Standard Deviation	Min.	Max
Home Buying Index	185	159.0757	15.48613	119	185

Summary of Home Buying Index 1992-2008 (Non-Homeowners)

Variable	Observations	Mean	Standard Deviation	Min.	Max
Home Buying Index	185	138.1767	22.2641	65.62963	176.7914

We then compared the homeowner and non-homeowner indexes for selling data:

Summary of Home Selling Index 1992-2008 (Homeowners)

Variable	Observations	Mean	Standard Deviation	Min.	Max

Home Selling Index	185	119.227	31.47657	18	164

Summary of Home Selling Index 1992-2008 (Non-Homeowners)

Variable	Observations	Mean	Standard Deviation	Min.	Max
Home Selling Index	185	106.0596	24.92768	22	161.2857

We found the same law at work in the selling data as we did in the buying data: homeowners are considerably more optimistic than non-homeowners (see Appendix Tables 1.B and 2.B for t-tests). This result was confirmed with paired t-tests comparing the means of each reason for homeowners and non-homeowners (see Appendix Tables 3.B and 4.B for t-tests). All good reason means were significantly higher for homeowners and all bad reason means were significantly higher for non-homeowners.

Also of note is the relative weight each group gives to low interest rates and low prices. For buying data, both rank interest rates as most important and then low prices, but the ratio of the mean response of these variables for homeowners is about 3:1, whereas it is only 2:1 for non-homeowners. This is shown below.

Homeowner Low Price and Low Interest Rate Response Variables

Variable	Observations	Mean	Standard Deviation	Min.	Max
(1) Good- Low Prices	185	14.93984	8.079826	5.04	40.6
(2) Good- Low Rates	185	43.61341	17.05729	10.88	73.08

Non-homeowner Low Price and Low Interest Rate Responses Variables

Variable	Observations	Mean	Standard Deviation	Min.	Max
(1) Good- Low Prices	185	11.42658	6.491393	-1.387778	31.64545
(2) Good- Low Rates	185	24.87025	13.87228	-25.61714	64.18079

Heterogeneity in the Saturated Model

Noticing these differences in sentiment across homeowners and non-homeowners prompted us to examine our saturated model seperately with regard to homeowners and non-homeowners. This yielded four regressions provided below: saturated models determining buying sentiment of homeowners and non-homeowners as well as saturated models determining selling sentiment of homeowners and non-homeowners.

Regression of Homeowner Buying Index on Contract Interest Rate, Nominal Median House Price in thousands of dollars, De-trended House Price in thousands of dollars, Inflation, Unemployment Level, Quarterly Percentage

Change in Contract Interest Rate, Yearly Percentage Change in Unemployment Level, and the Index of Consumer Expectations from 11/1992 to 3/2008:

Home Buying Index $= \beta_0 + \beta_1$ Contract Rate $+ \beta_2$ NominalHousePrice $+ \beta_3$ DetrendedHousePrice $+ \beta_4$ Inflation $+ \beta_5$Unemployment $+ \beta_6$ ContractRateChange $+ \beta_7$UnemploymentChange $+ \beta_8$ConsumerExpectations $+ u$
Observations = 173
$R^2 = .8565$

Independent Variables	Coefficient	Standard Error	T	P-Value
Contract Rate	-23.54465	1.457343	-16.16	0.000
Nominal House Price	-.5316172	.0323064	-16.46	0.000
De-trended House Price	.2790933	.0631927	4.42	0.000
Inflation	-2.594505	.8197897	-3.16	0.002
Unemployment	-3.656629	1.247736	-2.93	0.004
Contract Rate Change	44.38337	13.99986	3.17	0.002
Unemployment Change	20.94994	4.806432	4.36	0.000
Consumer Expectations	.3439705	.0738461	4.66	0.000
Constant	400.3327	22.95007	17.44	0.000

Regression of Non-Homeowner Buying Index on Contract Interest Rate, Nominal Median House Price in thousands of dollars, De-trended House Price in thousands of dollars, Inflation, Unemployment Level, Quarterly Percentage Change in Contract Interest Rate, Yearly Percentage Change in Unemployment Level, and the Index of Consumer Expectations from 11/1992 to 3/2008:

Home Buying Index $= \beta_0 + \beta_1$ Contract Rate $+ \beta_2$ NominalHousePrice $+ \beta_3$ DetrendedHousePrice $+ \beta_4$ Inflation $+ \beta_5$Unemployment $+ \beta_6$ ContractRateChange $+ \beta_7$UnemploymentChange $+ \beta_8$ConsumerExpectations $+ u$
Observations = 173
$R^2 = .7650$

Independent Variables	Coefficient	Standard Error	T	P-Value
Contract Rate	-18.74585	2.684573	-6.98	0.000
Nominal House Price	-.627704	.0595116	-10.55	0.000

De-trended House Price	.4254496	.1164072	-3.65	0.000
Inflation	-3.01995	1.510135	-2.00	0.047
Unemployment	-3.042572	2.298456	-1.32	0.187
Contract Rate Change	5.460214	25.78916	0.21	0.833
Unemployment Change	21.85534	8.853933	2.47	0.015
Consumer Expectations	.6204988	.1360319	4.56	0.000
Constant	336.0301	42.27635	7.95	0.000

Regression of Homeowner Selling Index on Contract Interest Rate, Nominal Median House Price in thousands of dollars, De-trended House Price in thousands of dollars, Inflation, Unemployment Level, Quarterly Percentage Change in Contract Interest Rate, Yearly Percentage Change in Unemployment Level, and the Index of Consumer Expectations from 11/1992 to 3/2008:

Home Selling Index $= \beta_0 + \beta_1$ Contract Rate $+ \beta_2$ NominalHousePrice $+ \beta_3$ DetrendedHousePrice $+ \beta_4$ Inflation $+ \beta_5$Unemployment $+ \beta_6$ ContractRateChange $+ \beta_7$UnemploymentChange $+ \beta_8$ConsumerExpectations $+ u$

Observations = 173

$R^2 = .8065$

Independent Variables	Coefficient	Standard Error	T	P-Value
Contract Rate	-44.1593	3.376404	-13.08	0.000
Nominal House Price	-.8950215	.0748482	-11.96	0.000
De-trended House Price	2.398863	.1464061	16.38	0.000
Inflation	13.10736	1.899307	6.90	0.000
Unemployment	-20.22267	2.890783	-7.00	0.000
Contract Rate Change	146.6437	32.43519	4.52	0.000
Unemployment Change	80.498848	11.13565	7.23	0.000
Consumer Expectations	1.594552	.1710882	9.32	0.000
Constant	498.8439	53.17123	9.38	0.000

Regression of Nonhomeowner Selling Index on Contract Interest Rate, Nominal Median House Price in thousands of dollars, De-trended House Price in thousands of dollars, Inflation, Unemployment Level, Quarterly Percentage Change in Contract Interest Rate, Yearly Percentage Change in Unemployment Level, and the Index of Consumer Expectations from 11/1992 to 3/2008:

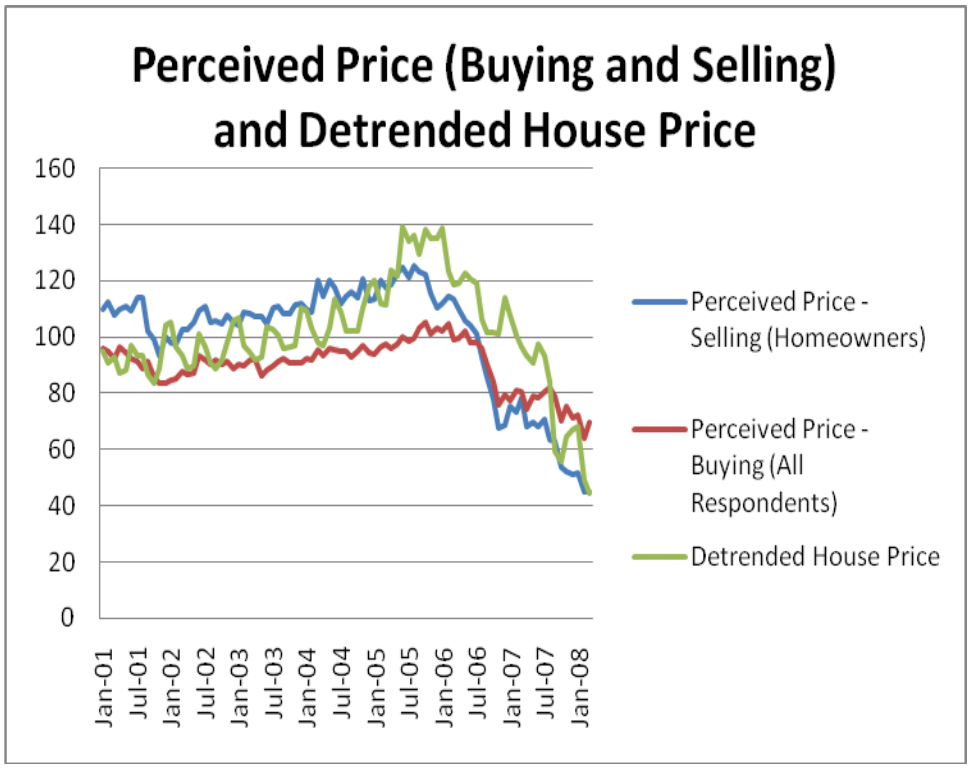
Home Selling Index $= \beta_0 + \beta_1$ Contract Rate $+ \beta_2$ NominalHousePrice $+ \beta_3$ DetrendedHousePrice $+ \beta_4$ Inflation $+ \beta_5$Unemployment $+ \beta_6$ ContractRateChange $+ \beta_7$UnemploymentChange $+ \beta_8$ConsumerExpectations $+ u$
Observations = 173
$R^2 = .6845$

Independent Variables	Coefficient	Standard Error	T	P-Value
Contract Rate	-23.09554	3.301614	-7.00	0.000
Nominal House Price	-.4274387	.0731902	-5.84	0.000
De-trended House Price	1.67112	.1431631	11.67	0.000
Inflation	6.72317	1.857236	3.62	0.000
Unemployment	-19.32654	2.82675	-6.84	0.000
Contract Rate Change	92.63923	31.71673	2.92	0.004
Unemployment Change	36.88916	10.88898	3.39	0.001
Consumer Expectations	.9992224	.1672985	5.97	0.000
Constant	329.905	51.99345	6.35	0.000

Interestingly, both homeowner saturated models have noticeably higher R^2 values than their non-homeowner counterparts. This indicates that more of the variation in homeowners' sentiment can be explained by the real-world macroeconomic variables we have identified as the most influential drivers of consumer sentiment. Furthermore, in both the buying and selling saturated models, nonhomeowners are less sensitive than homeowners to the contract interest rate as well as the quarterly change in contract interest rate. Since the homeowners, who are doing all of the selling, are more informed than the non-homeowners, those we assume to be doing a good part of the buying, changes in the macro-economy that might affect price will favor the sellers. In other words, if there is a disparity in perceived price between sellers and buyers, shifts in actual price will favor the sellers because they best adjust to the new information. So if there is indeed a disparity in sellers' and buyers' perceived price level, prices will tend to rise to the sellers' perceived price as opposed to lowering to the buyers' perceived price.

An Unsustainable Optimism

This disparity in perceived price between homeowners and nonhomeowners is confirmed in the data. More importantly, this disparity continues throughout the 4 year boom period—a reflection of unmitigated optimism on the part all agents in the housing market. When asked about buying, the general optimism tended to dictate a perception of low prices, while that same optimism for selling tended to dictate a perception of high prices. The prolonged disparity is illustrated in the graph below.



We decided to look at data starting in 2001, since it seemed that we never found any structural breaks or exceptional trends before that time. It is very obvious that from January 2001 all the way up to the housing crash in 2006, there was an unsustainable trend in that consumers were largely basing their buying sentiment on the perception of low prices while basing their selling sentiment on the perception of high prices. We hypothesize that this unmitigated optimism led to an unsustainable rate of increase in the price of housing since sellers (homeowners) are more informed regarding macro-economic indicators and thus shifts in price will favor them. In addition, those looking to buy homes continued to think that prices were low, and therefore that it was a good time to buy. This caused demand to shift outward and prices to rise sharply until 2006. At this point there was the realization that prices were far above fundamentals, and prices subsequently dropped.

Conclusion

This paper started as an exploration of the University of Michigan Consumer Sentiment data, specifically the questions about the housing market. As our analysis progressed we were able to report some interesting findings, test some economic theories, and propose a hypothesis on the housing boom and crash.

We notice from our summary statistics that since 1978 both buying sentiment and selling sentiment have been mainly positive. Although this would at first glance seem slightly odd, we propose that the housing market is not, in fact, a zero-sum game. Low interest rates, for one, benefit both the buyer and the seller. We present Inventory-Sales ratio as a way to reinforce that the housing market is not a zero-sum game.

Splitting the data into shorter periods of time, we found that there was a nonexistent or negative relationship between the two indexes. Over the longer 15 year period of time, our final saturated models for buying and selling sentiment are well-explained by the same economic indicators, all with the same signs (except for inflation). The similarity in signs may suggest that consumers base their sentiment of buying and selling conditions more on general economic wellbeing—when rates are low, unemployment is low, and expectations are good—than on any one particular reason specific to buying or selling. This finding helps explain why the buying and selling sentiment during the housing boom were simultaneously optimistic; an optimism that contributed (if not led) to a purely speculative housing bubble.

We also notice from our summary statistics that consumers mainly base their opinions on housing on interest rates and prices. However, as we report in Section II, consumers do not fully understand the main determinants of their own sentiment, exhibiting inflation illusion in prices and interest rates. With this misunderstanding in mind, we measure in Section III how much consumer sentiment is based on actual market fundamentals. We are able to explain about 80% of buying and selling sentiment with real-world variables that correspond to the reasons consumers are giving for their beliefs about the housing market.

In our last section we use the findings in the previous section in order to explain the housing market boom and subsequent crash. We believe that the asymmetry we find in beliefs about prices between homeowners and nonhomeowners created an unsustainable boom in prices. The crash was inevitable as the less-informed potential buyers finally realized that prices were well above fundamentals.