Reviewer's report

Title: Identification and quantification of change in Australian illicit drug markets

Version: Date: 27 November 2005

Reviewer: Jonathan Caulkins

Reviewer's report:

General

Background: The Australian drought or shortage is one of most important events in history for helping us understand drug markets. It has already engendered a considerable literature to which the authors of the current manuscript have made many and arguably the greatest contributions. So I think the question is not so much whether what this paper does is interesting and worthwhile but rather whether (1) is it done, presented, and interpreted in the clearest and most useful way and (2) does this paper make a sufficient marginal contribution above and beyond previous articles on the drought by these and other authors.

With respect to clarity of exposition, there are two sets of issues that ought to be easy to deal with a bit of rewriting (avoiding reference to “peaks” and “troughs” of some epidemic cycle and better presentation of the SiZer results).

With respect to marginal value of this paper: In some sense this paper shouldn’t be necessary. It is abundantly clear from other papers and the speed of the change in market-related indicators that the heroin markets in Australia experienced a supply side shock in late 2000 or January 2001. To be a bit impatient and undiplomatic, one would have to be ignorant and/or obtuse not to concede that point. Unfortunately, I have heard from some Australian colleagues that those might be pretty good descriptors of one not small faction in the row at the latest APSAD meeting (Australasian Professional Society on Alcohol and Other Drugs). My sense from afar is that there is a large block of drug ‘scholars’ who are so vehemently anti-enforcement that they cannot bring themselves to concede that a supply shock could ever produce a good outcome (such as the observed decline in overdose), for fear that the forces of repression will seize on that evidence to justify greater enforcement spending and incarceration. It is rather depressing that supposedly objective scientists could let their distaste for the potential policy implications of some data skew their interpretation of those data. Reductio ad absurdum should proceed from logical inconsistencies not from political inconveniences.

But, depressing or not, that is the world we live in, and the best antidote to obstinate refusal to admit what is plain in the data themselves may well be to marshal fancier statistical methods to “prove statistically” what ought to be evident at the surface. Hence, this paper, perhaps with some edits, may be useful and important.

Clarity of Exposition

One challenge for the manuscript’s authors is that the skeptics they seek to persuade to not have – as far as I can tell – a coherent story, so attacking it is like striking at a phantom. However, the authors should do some combination of (1) being clearer about the alternative theories they are trying to rebut and/or (2) de-emphasizing the broader theoretical arguments and just presenting the data results for what they are. The current text is a bit vague about the other camp’s views, and that is not a good situation.

The second paragraph presents one version of the alternative viewpoints: the decline was either a drought or the end of a glut. OK. That’s clear enough, but (1) other versions of the two camp’s perspectives dominate elsewhere in the text and there is no reason why there couldn’t be both an end of a glut and, in a addition, a discrete drought. My understanding is that both were in fact
occurring. There was a 20-month steady decline in heroin purity leading up to the precipitous collapse in purity in December 2000 or January 2001. The preceding 20 months might look like the tail end of a glut, but it is obvious that something changed abruptly in December 2000 or January 2001, and I see no reason not to call that precipitous change a “shortage”. (For that matter, I don’t get what wasn’t politically correct about the previous term of art, “drought”, but apparently it has fallen out of favor.)

Later in the second paragraph the text states “The question of whether a sudden collapse is a natural part of the final phase of an ‘epidemic’ cycle in illicit drug markets, or whether such a sudden collapse must be sparked by independent events …”

I spend quite a bit of time thinking about “epidemic cycles” of drug use and markets, and I am not aware of any serious model that claims that the end of the epidemic of a dependence-inducing substance such as heroin involves a market collapse. Indeed, I would expect most models to do the opposite, to have a slow tapering off. We may see a collapse in prevalence but not in demand and, hence, not in market equilibrium price, purity, or quantity, precisely because of the long-term stability of the habits of the minority of heavy users who dominate demand. This is documented for the US cocaine markets (cf. Everingham and Rydell’s 1994 RAND monograph which showed that cocaine demand was stable even as numbers of light users and, hence, total users, fell sharply because dependent users with low quit rates accounted for so much of the demand). Given career lengths of heroin use (cf. Anglin et al.’s work at UCLA among others), I would expect the same principle to apply to heroin markets. Furthermore, historically low purity in drug markets has been associated with a shortage of supply relative to demand. The end of a drug epidemic cycle connotes a decline in demand, so it is not obvious to me why even if there were an implausibly rapid decline in heroin demand that should lead to a decline in purity. One can always cook up convoluted stories. E.g., the market drops below critical sustainable mass, so that reversal of “enforcement swamping” tips the market back to a high-price, low-quantity equilibrium. However, my understanding of enforcement risks in Australia makes that hard to swallow in addition to being quite speculative. (Furthermore, the enforcement swamping dynamic should be strongest in local markets, so there is no obvious reason why the tipping point should have been reached at nearly the same time in each of the Australian states.)

Bottom line #1: I simply don’t understand what the authors’ “opponents” have in mind when they talk about a natural collapse at the end of an epidemic cycle.

Bottom line #2: If I’m right that you have to tell a pretty complicated story, then it’s not clear that the present data analysis would really refute such complicated stories. If I start out with a nonlinear endogenous epidemic model, and soup it up with enforcement swamping, tipping points, etc. to make it generate a “collapse at the end of an epidemic cycle” (whatever exactly that means), then presumably I’ve created a model that exhibits threshold behavior and could endogenously produce time series that show one thing for the 24 months before December 2000 and then show a precipitous decline in the next few months. Now this is a bit of a weird nit with the current paper because I doubt anyone has ever formally presented such a model in the first place, and any such model might be forced/contrived, but if the “opposition” is claiming there is some sort of natural collapse to a generally declining market than how does showing that a 24-month long special phase of the market is followed by a precipitous crash disprove their theory? (Same question arises from first sentence of second paragraph of the Discussion. It is unlikely that the shortage represents a natural correction after the end of a long period of decline since the principal component … is uncorrelated with the principal component representing long term trend. I don’t know. I think it’s unlikely that it was a natural correction since no one that I know of has a market model that predicts such a correction. As for the lack of correlation of the factors – isn’t that an automatic consequence of the way PCA works?)

Second (related) issue RE clarity of exposition
The discussion section and the earlier discussion of results fall into the trap of presuming that the 24 months preceding the drought can reasonably be called the “peak” of the epidemic cycle (and that the other times can be called a “trough”). Are you (current authors) truly committed to the idea that
there is such a thing as a peak and that it occurred at exactly that time? If so, then I’d like you to state precisely what you mean by that concept and prove to me that it is a fair description of what was going on. If not, then ditch that language or at least use it only at arm’s length. E.g., just say that “principal component #2 shows that the 24 months before the onset of the drought were somehow qualitatively different than were the other periods in our time series.” You might even say, “Some people believe that this period was a ‘peak’ of the epidemic. Even if that were true, it would be incorrect to view most of the variation in these time series as being attributable to the ending of that peak or glut since that component only explains 13% of the variance and principal component 1 explains three (or was it four) times as much of the variance.

I think the data show that the ending of whatever made the preceding 24 months special does not explain as much variance as does whatever happened right around January 2001. I see no reason not to call what happened around January 2001 a drought or a shortage. But I also see no reason to anoint what happened in the preceding 24 months with any special name or term, and I predict that even after we (the academic community collectively) figure out what was special about those 24 months, we won’t end up calling it the “peak of the epidemic” for any deep reason. (That is, those 24 months could in fact have witnessed the maximum of some indicator series, but that’s a simple ex post descriptive statement, not a statement that would have specific implications about what would have happened had there not been a drought.)

Aside: The debate about the cause of the shortage is not “academic” in the sense of being unimportant. It may be “academic” in the sense of being unresolved, but there have been drug shortages in the past (US heroin early 1970s, US cocaine September 1989 through 1990, US methamphetamine after pre-cursor controls, etc.), and there will be more such shortages in the future. Supply disruptions have effects, both good and bad, and they are the stated objective of many policy objectives. So they are anything but unimportant.

Technical Material

PCA

I have never seen PCA applied to time series data so I can’t comment on that as a methodologist. Hopefully you have some statisticians who are skilled in time series methods looking at the paper. Obviously the issue on my mind is serial autocorrelation of errors.

With regular PCA, one of the trickiest issues is interpretation of the factors. By analogy, why should I be sure the principal component #2 represents the difference between the “peak of the epidemic cycle” and “other times”? Couldn’t it just as easily represent the contrast between the tail end of a glut and other times? Those are very different concepts with different implications. Or at least I think they are. I think I know what the tail end of glut means. It means that there was a point at which too many factors of production were being supplied to the market relative to the long-run equilibrium return those factors could earn, so those factors were being withdrawn (or at least being injected at less than their rate of “depreciation”). If there are information lags, asymmetries, or one simply believes in linear adjustment dynamics instead of some Chicago School perfect foresight model, you can generate a story of a transitory market disequilibrium being eroded.

I’m less sure I know what the “peak of an epidemic cycle” means. Presumably it means different things for indicators of incidence, prevalence, demand, and for market clearing quantities. Such distinctions aren’t made. Indeed, the paper doesn’t even spend much time telling the reader what the 17 series are.

So, my suggestion is that the authors either get explicit about what the mean by their interpretation (define peak of an epidemic cycle) or back off from terms that seem to convey a lot of meaning but are actually ambiguous. E.g., they could call factor #2 the contrast between the period leading up to the shortage and other periods within the time series.

First sentence of the discussion calls the two principal components “contrasting information”. I
wouldn’t have picked the word “contrasting” for that sentence because it connotes that there is some conflict or inconsistency between them. It seems to me that the analysis suggests that there were two “pieces” of systematic variation in these time series, the first of which was time precisely at the traditional onset of the heroin shortage and the other pertains to the preceding 24 months. Those pieces are distinct and distinguishable, but not necessarily in tension.

I should add that it is interesting/exciting to have so many (17) different time series together. That certainly goes a long way toward assuaging concerns that the shortage was a statistical artifact of the idiosyncracies of one or two data indicator series.

SiZer
This is also the first time I’ve seen SiZer analysis. From what I understand of it, my main question is why I should expect a socio-economic system’s time series to have characteristic “frequencies” and, hence, to have signal concentrated in certain bandwidths. Those concepts make perfect sense to me for electronic circuits, but why does the variation in a socio-economic system have to have just one dominant frequency? Couldn’t the frequency for the “ending of the peak of an epidemic” (whatever that means it seems to pertain to demand shifts) be different than the frequency for a “drought” (which seems to be a supply shift). Perhaps the authors are claiming that the frequency pertains to the rate of adjustment for any market disequilbrium, but it’s not obvious to me why the speed of adjustment need be the same for shifts in demand as for shifts in supply or other shocks.

I may not be the only reader for whom SiZer is new, so the authors should work harder at clarifying the exposition, particularly the paragraphs above the place where Figure 5 will be inserted. Some detailed comments/questions.

The sentence beginning “In this plot regions …” This sentence says that both the increases and decreases are shown in blue whereas in fact decreases are show in red.

Next sentence. You say there is an increase then a decrease between months 0 and 35. First of all it’d be better to say increase (months 0 – 25) then decrease (months 25 – X). Second, that’s the right interpretation when the bandwidth is 2, but for other frequencies the decrease doesn’t end at month 35 but continues longer.

Note: I don’t think you need to have the red regions be separated by a plateau to make your point. I think your main point is that the rate of decline (slope) is significantly steeper after January 2001 than it was before January 2001, not that the slope was level before January 2001.

“The bandwidth of this best fit is approximately 2 …” Tell reader in the text that log10(2) = 0.3 so they know what part of the Figure to focus on without pulling out their calculator.

“Larger bandwidths …” So it seems to me that shorter bandwidths, including the best fit bandwidth, detect a ‘double dip’ pattern, whereas longer bandwidths that smear more together see only one broad decline. Is that a reasonable way to think about it?

Aside: I’m not clear how discriminating the “best fit” estimator is for bandwidth. Sure 2 is the best fit, but are 3 or 10 almost as good, or are they impossibly inconsistent with the data so I should really only pay attention to a bandwidth of 2?

Related aside: Even the text seems to waffle on whether to pay attention to the other bandwidths. The second to the last paragraph of the paper touts a result’s holding over all bandwidths whereas this portion seems to suggest that focusing on the best fit bandwidth is the key. For that matter, is the vertical range in Figures 5 and 6 appropriate? Is it meaningful to display the results for bandwidths less than one? For bandwidths as large as $10^2 = 100$? Maybe the figure should cover less range and be expanded???
What is the faint dotted line in Figures 5 and 6?

Figure 6 text talks about a best fit bandwidth having an increase between months 5 and 12 and another increase from 45 to 51, period. But doesn’t it also show a decrease after month 60? A few sentences later text talks about the increase never reversing, but for bandwidths slightly above the best fit bandwidth there is a red rectangle around months 25-31. Are you telling me that the microscopic difference in bandwidths between the best fit and that red rectangle is too large to view the red rectangle as meaningful? If so, why do you show anything except the best fit bandwidth. Furthermore, it seems to me that purple doesn’t mean level so much as inability to reject a null of level, so how do you conclude positively that the two periods of increase are unbroken? Just looking at the family overlay it sure looks like they were broken. Last para on Figure 6 talks about a sudden change between July 2000 and February 2001. Is it fair to call a change over six months sudden? I suspect the drought is actually a lot more sudden than six months, but the nature of the SiZer smoothing process actually spreads out some of the change, diluting the case that the authors want to make that the change was sudden.

-------------------------------------------------------------------------------
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

-------------------------------------------------------------------------------
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

-------------------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)

Minor:
Figure 1 would be improved if there we just one panel showing all 17 plots but with the top two plots scaled so that displaying them does not compress the other 15 down along the x-axis. A wider figure should make it look a bit less jumbled.

Or, if you were going to use two panels, consider lumping together like measures in terms of how one would expect them to behave. E.g., the six or seven measures of substitution into stimulants on one panel, the five measures of acquisitive crime/prostitution in another, and the three measures of opiate use in another.

(A perennial complaint with analyzing time series related to the Australian drought is that there is no a priori theoretical reason to expect them to respond similarly to a given market shock. E.g., a supply shock may increase treatment demand in the short-run but reduce it in the long-run, making it look like things got worse at first and then improved, whereas focusing on price or purity one might see the exact opposite (initial market effect and then a return toward long-run equilibrium). If the elasticity of demand is -1, then a supply shock may have no effect on prostitution or acquisitive crime. Presumably the short-term elasticity is smaller, in absolute value, than the long-term elasticity. Depending on which if either is larger in absolute value than 1, one could tell various stories about the direction of effect in the short and long run. There is also no reason that the duration of the “short run” need be the same with respect to acquisitive crime and treatment demand. One can even tell stories about how large shifts in drug use and associated problems can feed back into the allocation of enforcement resources to drugs as opposed to other activities. In short, I wouldn’t expect all seventeen series to parallel each other even if there were a market shock.)

Figure 2 – for similar reasons – since you don’t identify the separate lines or intend the readers to be able to figure out which is which, consider using colors to lump the series together as just described. E.g., use red for all the series reflecting criminal income generation, blue for the stimulant related
series, and green for the opiate use related series.

**What next?:** Accept after minor essential revisions

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes

**Declaration of competing interests:**

possibly yes RE non-financial competing interests. See note to editor above.