Mon. Not. R. astr. Soc. (1990) **242,** *Short Communication,* 43P–47P

Galaxy correlations on large scales

S. J. Maddox,[1] G. Efstathiou,[1] W. J. Sutherland[1,2] and J. Loveday[1,2]

[1] *Department of Astrophysics, Keble Road, Oxford OX1 3RH*

[2] *Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*

Accepted 1989 October 23. Received 1989 September 7; in original form 1989 June 29

SUMMARY

We present our first results on large-scale structure in the Universe from a uniform survey of $\gtrsim 2$ million galaxies brighter than $b_J = 20.5$ constructed from machine scans of 185 UK Schmidt plates. We show that over a range of three magnitudes the galaxy two-point angular correlation function, $w(\theta)$, scales with depth as expected if we are measuring real clustering in the galaxy distribution. Our correlation functions show a break from a power law at roughly the same physical separation as found by Groth & Peebles from their analysis of the Lick catalogue, but our measurements decline much more gently from a power law on larger scales. We argue that Groth & Peebles may have removed some intrinsic clustering when they corrected for large-scale gradients in the Lick counts. Our analysis has important implications for theories of the formation of large-scale structure. In particular, our results imply more large-scale clustering than predicted by popular versions of the Cold Dark Matter cosmogony.

1 INTRODUCTION

The galaxy two-point correlation function, $\xi_g(r)$, deserves a special place amongst statistics of the galaxy distribution. On scales $\gtrsim 10h^{-1}$ Mpc (where h is Hubble's constant in units of 100 km s^{-1} Mpc^{-1}) correlations between galaxies are weak [$\xi_g(r) \ll 1$] so there is a reasonable expectation that $\xi_g(r)$ can be related by linear perturbation theory to fluctuations in the early Universe. Furthermore, most models in which perturbations are generated during an inflationary phase in the early Universe (see e.g. Bardeen *et al.* 1983) predict fluctuations that obey Gaussian statistics. If these models are correct, *all* aspects of linear structure can be described by the two-point correlation function. Thus, if we knew $\xi_g(r)$ on large scales, we would have an extremely powerful constraint on models of the early Universe and, if the fluctuations are Gaussian, we would have a complete description of large-scale structure.

However, as many investigators have found, it is extremely difficult to measure $\xi_g(r)$ on scales $r \gtrsim 10h^{-1}$ Mpc. Several years ago, we began a programme to produce a deep, uniform catalogue of galaxies over much of the southern sky from machine scans of UK Schmidt plates. The final catalogue samples a huge volume of space ($\sim 10^8 h^{-3}$ Mpc^3), much larger than the volume sampled in any redshift survey. Of course, from a two-dimensional catalogue we have to infer $\xi_g(r)$ from measurements of the angular two-point correlation function $w(\theta)$. This approach has been followed by Groth & Peebles (1977, hereafter GP77) who measured $w(\theta)$ from the Lick survey (Shane & Wirtanen 1967; Seldner *et al.* 1977, hereafter SSGP). However, our catalogue is more accurate than the Lick survey and, as we demonstrate below, is less subject to spurious clustering.

We summarize the catalogue in Section 2 and present estimates of $w(\theta)$ in Section 3. Our results are discussed in Section 4, where we compare with $w(\theta)$ from the Lick survey (GP77) and with the predictions of the Cold Dark Matter (CDM) theory (e.g. Davis *et al.* 1985; White *et al.* 1987).

2 THE APM GALAXY SURVEY

The amplitude of $w(\theta)$ is small at large angular scales, hence it is vital that a two-dimensional catalogue be uniform over the sky. The requirements of high accuracy and large sky coverage can be achieved only by using an automatic plate scanning machine. We have used the SERC Automatic Plate Measuring (APM) system in Cambridge (Kibblewhite *et al.* 1984) to digitize 5.8° × 5.8° from each of 185 copies of UK Schmidt J survey plates. Images are detected and analysed in real time, producing the position, magnitude, second moments and surface brightness profile for each object. The plates cover a contiguous area of 4300 square degrees, in the region $\delta < -20°$ and $b \lesssim 40°$.

There are 10 million images brighter than $b_J = 20.5$ in our survey, of which only 2 million are galaxies. The APM image parameters are used to distinguish between galaxies, stars and merged objects (mostly stars). We have visually checked more than 4000 images on several UKST and AAT plates and find that the galaxy sample is 95 per cent complete in the range $17 < b_J < 20.5$. Contamination from merged images is <6 per cent in the range $17 < b_J < 20$ but rises to 12 per cent at $b_J = 20.5$ which is close to the limit for accurate image classification. Adjacent plate centres are separated by 5°, so neighbouring plates have large overlaps enabling stringent control of systematic errors. We match galaxies in these overlaps and compute a magnitude correction for each plate so that residual errors in the overlaps are minimized (see SSGP, who apply a similar matching algorithm to the Lick counts). After matching, the rms of the overlap residuals is 0.06 mag, implying that the rms error in plate zero-points is 0.024 mag. We have included a correction for galactic obscuration $[A_b = C(\csc b - 1),\ C = 0.1]$ but this has very little effect on our results since we have restricted much of the present analysis to $|b| > 50°$. The correction for obscuration is uncertain, but it is unlikely that the coefficient C is larger than 0.2 (see e.g. Sandage & Tammann 1981, and references therein). The survey has been calibrated using photometric sequences of faint galaxies in 40 CCD fields distributed over the survey area.

Full details of our analysis techniques, performance of the APM machine, and comparisons between copy plates and originals are presented elsewhere (Maddox 1988; Maddox *et al.* 1990, in preparation). Plate 1 shows a map of the galaxy distribution in our survey over the magnitude range $17 < b_J < 20.5$.

3 THE TWO-POINT ANGULAR CORRELATION FUNCTION

We use two estimators to measure $w(\theta)$. On scales $\theta < 0.5°$, we use the coordinates of each galaxy to count pairs; our estimate of $w(\theta)$ is

$$w(\theta) = F\frac{DD}{DR} - 1, \qquad (1)$$

where DD is the number of data pairs with separations in the range $\theta \pm \delta\theta$. This is normalized by the term DR which is the pair count determined by cross correlating the data points with a set of points distributed at random within the boundary with F times the mean density of galaxies.

On angular scales $0.1° < \theta < 30°$ we compute $w(\theta)$ from maps of the galaxy distribution in an equal area projection. We apply the estimator

$$w(\theta) = \frac{\langle N_i N_j \rangle}{\langle N_i \rangle \langle N_j \rangle} - 1, \qquad (2)$$

where N_i and N_j are the counts in cells i and j, and the angular brackets denote an average over pairs of cells separated by an angle $\theta \pm \delta\theta$.

These estimates must be corrected for the dilution of clustering caused by contaminating merged images which we assume are uncorrelated. Thus, we multiply our estimates by $1/(1-f)^2$, where f is the residual contaminating fraction estimated from our visual checks.

Fig. 1 shows the mean of $w(\theta)$ for galaxies with $17 < b_J < 20$ in each of four separate zones containing 45, 45, 46 and 49 plates (corresponding roughly to equally spaced strips in right ascension). The error bars are computed from the scatter between the zones. This shows that $w(\theta)$ can be approximated by a power law at small angles; a least-squares fit to $w(\theta)$ between $\theta = 0.01°$ and $\theta = 1°$ gives a slope of -0.668 and an amplitude of 2.84×10^{-2} at $\theta = 1°$. At large angles the slope of $w(\theta)$ gradually steepens. The break at ~ 3° is extremely significant given the scatter between zones. The inset in Fig. 1 shows that $w(\theta)$ is very close to zero (to within $\sim 5 \times 10^{-4}$) for $\theta \gtrsim 6°$.

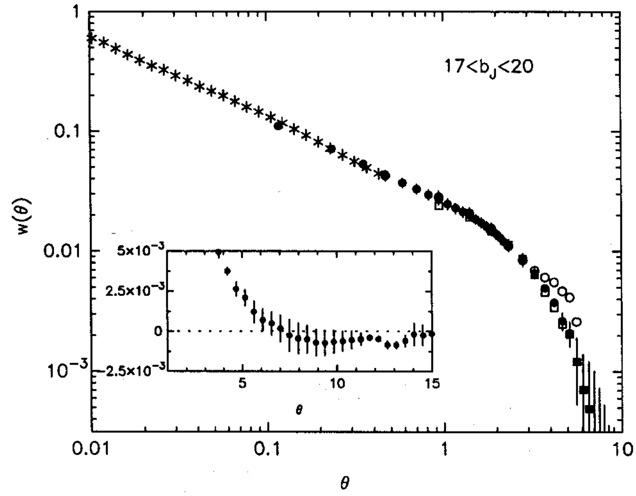

Figure 1. The angular correlation function $w(\theta)$ for galaxies in the magnitude range $17 \le b_J \le 20$. The crosses and filled circles show the mean $w(\theta)$ computed from equations (1) and (2), respectively for four zones each containing about 45 plates. The error bars show the standard deviation of the mean. Open circles and open squares show the mean of $w(\theta)$ for intra- and inter-plate pairs, respectively. The inset shows the mean $w(\theta)$ to large angles on a linear scale.

The open circles in Fig. 1 show intra-plate estimates of $w(\theta)$ in which we have used pairs of cells only if they are on the same plate. The open squares show inter-plate estimates of $w(\theta)$, in which we count pairs of cells only if they are on neighbouring plates. The differences between intra- and inter-plate estimates of $w(\theta)$ provide a measure of the variance of the plate matching errors (GP77). We find a mean offset of 1.7×10^{-3}, which implies an rms error in true number density on each plate of ~ 4 per cent. The measured value is in good agreement with the expected offset (1.2×10^{-3}) from photometric and classification errors in our final map.

A further test of our catalogue is shown in Fig. 2 where we present $w(\theta)$ for each of six disjoint 0.5 mag slices between $b_J = 17$ and 20.5. We have limited this analysis to the central 120 plates of our survey which corresponds to $b \lesssim -50°$. At such high $|b|$ no significant errors are introduced by uncertainties in the correction for galactic absorption or increased contamination from merged stellar images at low galactic latitudes.

We have scaled these estimates of $w(\theta)$ to the depth of the Lick catalogue using the relativistic version of Limber's equation (GP77, Phillipps *et al.* 1978) with $q_0 = 0.5$. The selection function for our survey is determined using the following model for the galaxy luminosity function as a function of redshift z:

$$\varphi(x)\,dx = \varphi^* x^\alpha \exp(-x)\,dx,\ x \equiv 10^{0.2(M^*_{b_J} - M)},$$
$$M^*_{b_J}(z) = M^*_0 + M^*_1 z,\ \alpha(z) = \alpha_0 + \alpha_1 z \qquad (3)$$
$$\varphi^* = 1.3 \times 10^{-2} h^3\,\mathrm{Mpc}^{-3},\ M^*_0 = -19.8,\ M^*_1 = 1,\ \alpha_0 = -1,$$
$$\alpha_1 = -2.$$

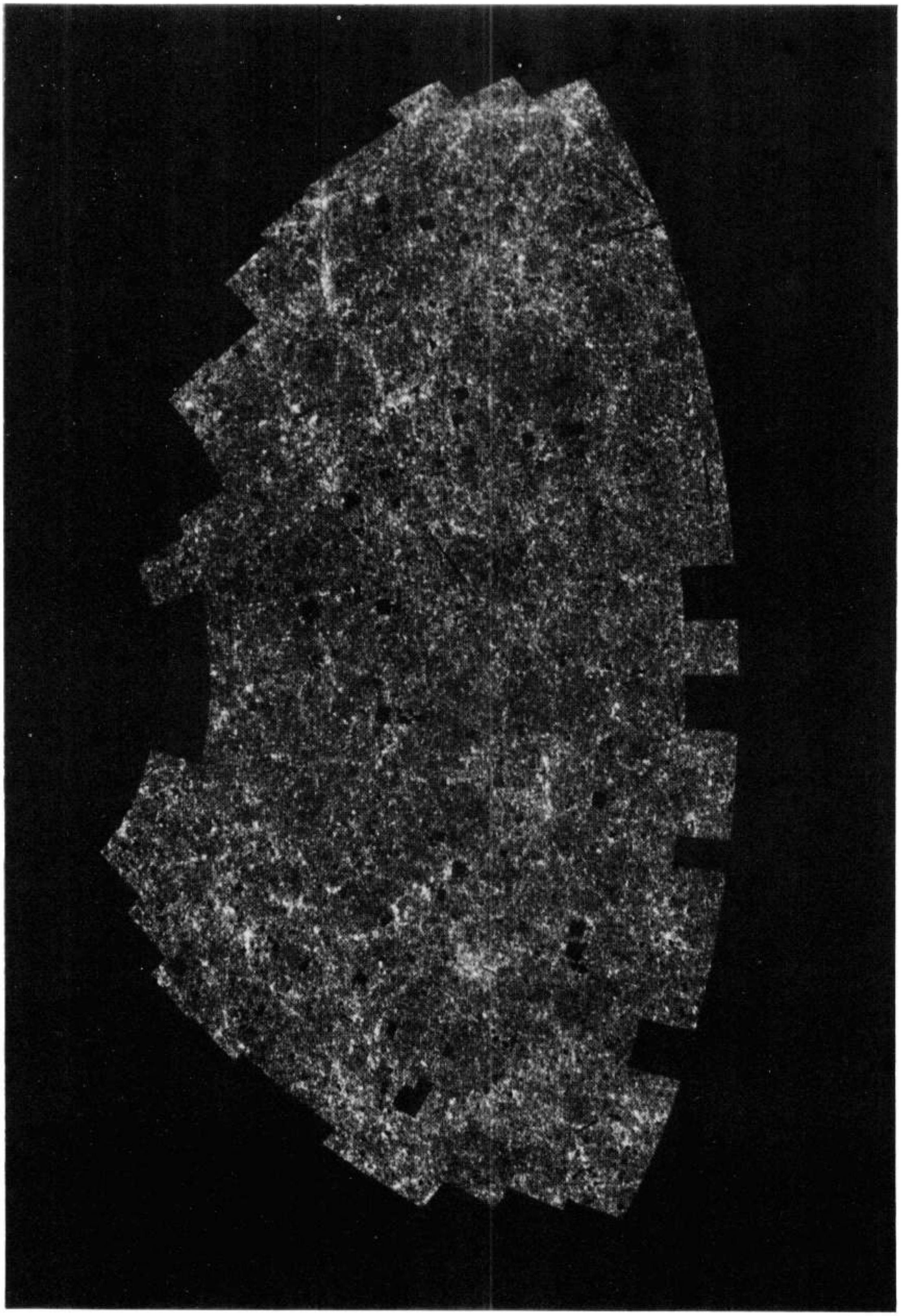

Plate 1. The distribution of galaxies with $17 \leq b_J \leq 20.5$ shown in an equal area projection centred on the southern galactic pole. The small empty patches in the map are regions that we have excluded around bright stars, nearby dwarf galaxies, globular clusters and step wedges.

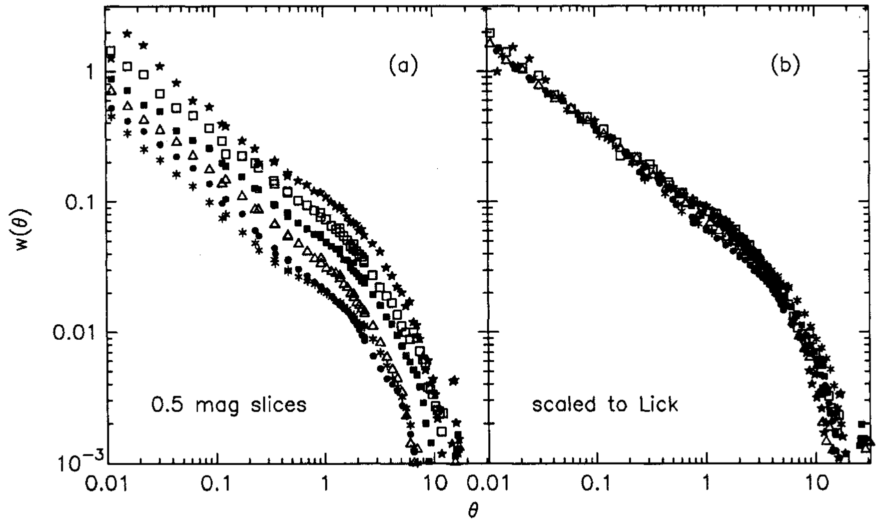

Figure 2. (a) Shows angular correlation functions for six 0.5 mag slices in the range $17.5 \le b_J \le 20.5$. (b) Shows the results from (a) scaled to the depth of the Lick survey as described in the text.

This parametric form is intended to model the evolution of the galaxy luminosity function (including the *k*-correction) and to be consistent with the following observations: (1) the low-redshift field galaxy luminosity function (Efstathiou, Ellis & Peterson 1988); (2) the galaxy number counts from our survey and from Tyson (1988); (3) the redshift distribution in the deep spectroscopic survey of Broadhurst, Ellis & Shanks 1988), and (4) the evolution of the luminosity function determined from Broadhurst *et al.*'s survey (Maddox *et al.*, in preparation).

We use (3) in Limber's equation and a two-power-law model for $\xi(r)$ ($\xi \propto r^{-1.66}$ for $r < r_1$, $\xi \propto r^{-3}$ for $r > r_1$) to compute the shifts in $\log(w)$ and $\log(\theta)$ which scale our results to the depth of the Lick survey (see GP77, fig. 15). Our model for φ produces the observed Lick galaxy surface density at a magnitude limit of $b_J = 18.4$. We also assume that the clustering pattern is stable in proper coordinates, i.e. $\xi(r, z) \propto 1/(1+z)^3$.

Fig. 2(b) shows the curves from Fig. 2(a) scaled to the Lick depth as described above. The agreement between these curves is excellent and the scatter is consistent with the errors expected from our plate-matching procedure. We expect our estimates to be too high by $\sim 2 \times 10^{-3}$ as a result of large-scale gradients in the plate-matching procedure (Section 4). This probably explains why $w(\theta)$ for the faintest magnitude slice is slightly high on scales $\gtrsim 1°$. Subtracting 2×10^{-3} from the estimates would improve the agreement between the scaled correlation functions shown in Fig. 2(b). The estimates for the brightest magnitude slice may be biased low by $\sim 2 \times 10^{-3}$ as a result of the finite size of the APM survey (see GP77, equation 29). Both of these effects are small and are of marginal importance to the results shown in Fig. 2. The scaling test thus provides important evidence that our angular correlation functions are measuring real clustering in the galaxy distribution rather than clustering induced by systematic errors.

4 DISCUSSION

In Fig. 3, we compare the scaled results in Fig. 2(b) with $w(\theta)$ from the Lick survey (GP77). Our results are in good agreement on scales up to $\theta \approx 3°$ but disagree on larger scales. GP77 find that $w(\theta)$ breaks sharply from a power law on scales of $\sim 3°$, whereas we find a more gentle decline. However, GP77's estimate of $w(\theta)$ is based on filtered versions of the Lick map and it is possible that filtering has removed a component of intrinsic clustering.

Groth & Peebles (1986a, hereafter GP86) show that filtering is required to remove artificial large-scale gradients in the Lick map. SSGP find that the mean square error in the fractional count in a single overlap strip after applying the Lick plate correction factors is $\sigma^2 = 0.025$. These plate-matching errors propagate through the network of plates and introduce large-scale gradients in the corrected counts, producing

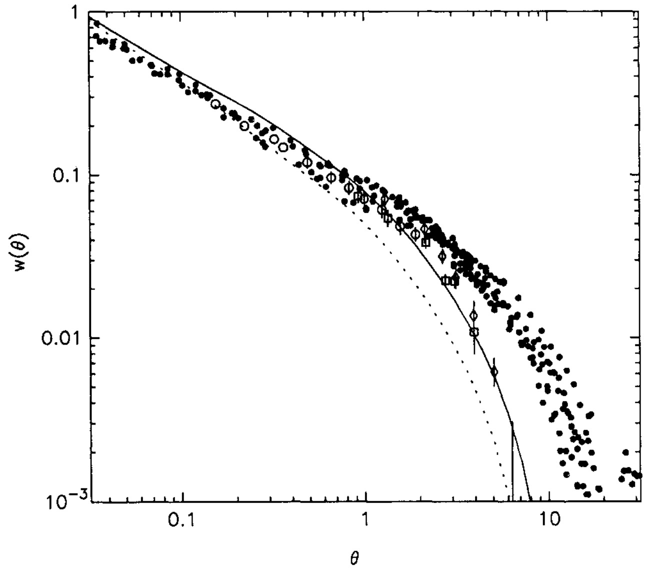

Figure 3. The filled circles show our estimates of $w(\theta)$ in six 0.5 mag slices (Fig. 2b) scaled to the Lick depth. The open symbols show $w(\theta)$ for the Lick catalogue from fig. 5 of Groth & Peebles (1977); the symbols have the same meaning as in their figure. The dotted and solid lines show computations of $w(\theta)$ based on the CDM model with $h = 0.5$ and $h = 0.4$, respectively, as discussed in the text.

a variance in the counts on a single plate of $\beta_{ii}^2 = 0.0127$. Subtracting the contribution from plate-to-plate errors ($4\sigma^2/25$), GP86 conclude that large-scale gradients introduce a variance of 0.0087. However, the polynomial map which GP86 fit to the data has a variance of 0.0194, more than twice the predicted value.[*] The difference between these two numbers could be caused by galactic obscuration, systematic errors in star–galaxy separation, or real clustering in the galaxy distribution. If we add 0.01 to GP77's determination of $w(\theta)$, we find excellent agreement with our estimates (Fig. 3) out to scales of $\theta = 10°$; even if we add only 0.005 to the GP77 results they would be compatible with ours within the errors. It is therefore plausible that Groth & Peebles have over-filtered the Lick data. Other possible sources of error in the Lick catalogue have been discussed by Geller, de Lapparent & Kurtz (1984), de Lapparent, Kurtz & Geller (1986) and Groth & Peebles (1986a,b). For the APM survey, we estimate that magnitude and classification errors give $\sigma^2 = 0.005$. Approximating our survey by a 13 × 13 plate network, the GP86 model gives $\beta_{ii}^2 = 2 \times 10^{-3}$, thus the effects of large-scale gradients should be of marginal importance even in our faintest magnitude slice (*cf.* Section 3). Because of the high accuracy of our catalogue, we have not had to filter our maps. A full discussion of the effects of plate-matching errors on our results will be presented elsewhere (Maddox *et al.*, in preparation).

Some spurious large-scale clustering could be introduced into our estimates from variable obscuration. To produce an offset of 0.005 in our estimates of $w(\theta)$ would require an absorption coefficient in the csc b model of $C = 0.43$ which is much too large and incompatible with the latitude dependence of the galaxy density in the APM survey. Patchy obscuration could, in principle, contribute to $w(\theta)$. However, $w(\theta)$ for the faintest magnitude slice provides an upper limit on such effects; even if the entire $w(\theta)$ for the faintest magnitude slice were caused by patchy obscuration, it would have little effect on $w(\theta)$ for the brightest three slices. It is therefore likely that patchy obscuration is unimportant.

In a previous article we estimated $w(\theta)$ from a small region of about 25 plates from the APM survey (Maddox, Efstathiou & Loveday 1988) and found good agreement with GP77's results. This does not conflict with the results shown in Figs 1 and 2, but is a consequence of the large sampling fluctuations in a 25 plate patch at these magnitude limits. Collins, Heydon-Dumbleton & MacGillivray (1989) have estimated $w(\theta)$ from a contiguous area of 24 Schmidt plates to $b_J = 20.5$. However, our analysis (*cf.* Fig. 1) implies that their region is too small for an accurate determination of $w(\theta)$ at $\theta \gtrsim 1°$.

The two curves in Fig. 3 show predictions for $w(\theta)$ based on the 'standard' biased Cold Dark Matter (CDM) model of Davis *et al.* (1985). We have computed $\xi(r)$ from the numerical simulations described by White *et al.* (1987) using their prescription for biasing; on scales $< 1h^{-1}$ Mpc we have extrapolated $\xi(r)$ as a power-law $\xi(r) = (r_0/r)^{-1.7}$, $r_0 = 5.7h^{-1}$ Mpc, since the numerical simulations have limited resolution on small scales. The dotted lines show results for $h = 0.5$; these provide a good fit to $w(\theta)$ on small scales but fail to match the Lick $w(\theta)$ on scales $\gtrsim 1°$ (see also Bond & Couchman 1988, for similar theoretical predictions). The solid line shows the result of lowering the Hubble constant to $h = 0.4$, while retaining all the other parameters fixed. These results match the Lick data on scales $\gtrsim 1°$ but predict too high an amplitude on small scales. However, neither model matches the results from the APM Galaxy Survey. Our survey provides strong evidence for large-scale power in the galaxy distribution which is difficult to reconcile with the standard CDM model of White *et al.* (1987). We emphasize, however, that the standard CDM model represents just one variant of the cold dark matter cosmogony and that it may be possible to construct variants which are compatible with our measurements. We will discuss some possibilities in a separate paper.

The detection of small-amplitude clustering on large scales is extremely sensitive to systematic errors. The tests that we have applied to the APM Survey indicate that we are measuring real galaxy clustering, but the results clearly need to be confirmed by other surveys. The sparse-sampled redshift surveys of bright APM galaxies (Loveday, Efstathiou & Peterson, in preparation) and IRAS galaxies (Lawrence *et al.*, in preparation), and the galaxy catalogue of Heydon-Dumbleton, Collins & MacGillivray (1989) are three ongoing projects that should provide important checks of our results.

ACKNOWLEDGMENTS

We are extremely grateful for all the assistance provided by the APM group (M. J. Irwin, P. Bunclark, M. Bridgeland) over the last few years and for numerous discussions with members of the Institute of Astronomy, Cambridge. We are also grateful to the University of Oxford for their support at a critical stage in the project. W. J. Sutherland and J. Loveday acknowledge the receipt of SERC studentships. S. J. Maddox has held an SERC studentship and research assistantship during the course of this work. We would like to thank Ed Groth and Jim Peebles for their comments on this paper.

REFERENCES

Bardeen, J. M., Steinhardt, P. J. & Turner, M. S., 1983. *Phys. Rev. D*, **28**, 679.

Bond, J. R. & Couchman, H., 1988. In: *Proc. Second Canadian Conference on General Relativity and Relativistic Astrophysics*, p. 385, eds Coly, A. & Dyer, C., World Scientific, Singapore.

Broadhurst, T. J., Ellis, R. S. & Shanks, T., 1988. *Mon. Not. R. astr. Soc.*, **235**, 827.

Collins, C. A., Heydon-Dumbleton, N. H. & MacGillivray, H. T., 1989. *Mon. Not. R. astr. Soc.*, **236**, 7P.

Davis, M., Efstathiou, G., Frenk, C. S. & White, S. D. M., 1985. *Astrophys. J.*, **292**, 371.

de Lapparent, V., Kurtz, M. J. & Geller, M. J., 1986. *Astrophys. J.*, **304**, 585.

Efstathiou, G., Ellis, R. S. & Peterson, B. A., 1988. *Mon. Not. R. astr. Soc.*, **232**, 431.

Geller, M. J., de Lapparent, V. & Kurtz, M. J., 1984. *Astrophys. J.*, **287**, L55.

Groth, E. J. & Peebles, P. J. E., 1977. *Astrophys. J.*, **217**, 385 (GP77).

Groth, E. J. & Peebles, P. J. E., 1986a. *Astrophys. J.*, **310**, 507 (GP86).

Groth, E. J. & Peebles, P. J.E., 1986b. *strophys. J.*, **310**, 499.

---

[*] We have simulated the GP86 model and find that an 18 × 18 plate network, the variance of β^2_{ii} from the model prediction is only 16 per cent; the model predictions can therefore be compared directly to the Lick map.

Heydon-Dumbleton, N. H., Collins, C. A. & MacGillivray, H. T., 1989. *Mon. Not. R. astr. Soc.,* **238,** 379.

Kibblewhite, E. J., Bridgeland, M. T., Bunclark, P. & Irwin, M. J., 1984. *Astronomical Microdensity Conference,* NASA Conf. Pub., **2317,** p. 277.

Maddox, S. J., 1988. *PhD thesis,* University of Cambridge.

Maddox, S. J., Efstathiou, G. & Loveday, J., 1988. *Large Scale Structures of the Universe, IAU Symp. No. 130,* p. 151, eds Audouze, J., Pelletan, M.-C. & Szalay, A., Kluwer, Dordrecht.

Maddox, S. J., Sutherland, W. J., Efstathiou, G. & Loveday, J., 1990. *Mon. Not. R. astr. Soc.,* in press.

Phillipps, S., Fong, R., Ellis, R. S., Fall, S. M. & MacGillivray, H. T., 1978. *Mon. Not. R. astr. Soc.,* **182,** 673.

Sandage, A. & Tammann, G. A., 1981. *A Revised Shapley-Ames Catalog of Bright Galaxies,* Carnegie Institution of Washington Publication 635.

Seldner, M., Seibers, B., Groth, E. J. & Peebles, P. J. E., 1977. *Astr. J.,* **82,** 249 (SSGP).

Shane, C. D. & Wirtanen, C. A., 1967. *Publs Lick Obs.,* **22,** Part 1.

Tyson, J. A., 1988. *Astr. J.,* **96,** 1.

White, S. D. M., Frenk, C. S., Davis, M. & Efstathiou, G., 1987. *Astr. J.,* **313,** 505.